No Kids, No Tech: How Shortages of Young Workers Hinder Firm Technology Adoption

Cäcilia Lipowski*

[Link to latest version]

January 8, 2024

Abstract

Firms in developed countries increasingly report shortages of skilled workers. This paper studies how shortages of young labor market entrants, particularly trainees, affect firm technology adoption. I exploit exogenous variation in trainee supply induced by an education reform in Germany in 2001. Based on a large firm panel survey and social security records, I show that a reduction in trainee supply decreases firm technology investments. This effect is explained by trainees excelling in learning new tech skills, provoking high capital adjustment costs when trainees are scarce. These findings dampen hopes of counteracting labor shortages by substituting labor with capital.

Keywords: Labor Shortages, Firm Investments, Endogenous Technological Change, Capital Adjustment Costs, Vintage-Specific Technical Skills

JEL: D22, D24, J10, J21, J24, O33

*ZEW Mannheim, caecilia.lipowski@zew.de

I am greatly indebted to Anna Salomons and Ulrich Zierahn-Weilage for their advice and support. For helpful comments and conversations, I thank Daron Acemoglu, Melanie Arntz, David Autor, Eduard Brüll, Katja Görlitz, Maarten Goos, Simon Jäger, Harald Pfeifer, Pascual Restrepo, Johannes Schmieder, Anna Waldman-Brown, and Nicolas Ziebarth. I thank conference and seminar participants at the EEA (Barcelona), EALE (Prague), TPRI (Boston), Skills-for-the-Future Conference (LISER) and at internal seminars at IZA, University of Mannheim, Utrecht University and ZEW Mannheim. The project was financially supported by the Leibniz Association through the Leibniz Professorship for Applied Labor Economics at the University of Heidelberg (P56/2017) and ZEW Mannheim.

1 Introduction

Firms in developed countries increasingly suffer from shortages of skilled labor, which are expected to further intensify due to demographic change (Lightcast, 2021; OECD, 2023). Labor shortages may have significant consequences for economic growth if, in response, firms adjust their investment behavior and technology adoption. However, the effect of labor shortages on firm investments remains largely unexplored and ambiguous. On the one hand, firms could respond by adopting labor-replacing technologies to compensate for the lack of workers. On the other hand, labor shortages may hinder the implementation of technologies that require worker skills. Identifying the causal effect of labor supply shortages on firm investments is challenging because labor supply reductions tend to evolve gradually; usually go hand in hand with changes in labor demand; and are often confounded by unobserved factors at the region, industry, or firm level.

In this paper, I overcome this identification issue exploiting an education reform and provide empirical evidence on the causal effect of supply-driven labor shortages on firm technology investments. I focus on shortages of young labor market entrants, i.e. trainees, because their availability may be the bottleneck to the adoption of technologies requiring up-to-date skills if trainees have an advantage in learning new skills compared to incumbent workers.

My identification strategy exploits a natural experiment created by an education reform. In 2001, two out of six East German federal states, henceforth "treated states", permanently increased the length of schooling required for the university entrance degree by one year.[1] In treated states, the reform caused a missing school graduation cohort from the upper school track, temporarily reducing the supply of young labor market entrants, while there was no comparable reduction in the other four Est German states, henceforth "control states". Since labor market entry in Germany is commonly via vocational training, the missing school graduation cohort translates into a missing trainee entry cohort and significantly reduces the stock of trainees in subsequent years. The reform-induced variation in trainee supply across time and states is plausibly exogenous to firms, especially since trainees are highly immobile.[2] The missing trainees from the upper school track can be thought of as currently unskilled but future middle-skilled professionals undergoing on-the-job training. They often work in white-collar occupations such as media, retail, or financial service occupations, which commonly require bachelor's or associate degrees in other countries like the US.

The German vocational training system provides an exceptional opportunity for studying implications of shortages of young labor market entrants. First, vocational training is omnipresent in the German labor market with two thirds of the workforce holding a vocational

[1] Büttner & Thomsen (2015); Morin (2015); Muehlemann et al. (2022); Marcus & Zambre (2019) and Dorner & Görlitz (2020) also exploit this and a similar reform affecting required years of schooling, studying the effect on grades, university enrollment, and trainee employment and wages.

[2] Only 2.2% of trainees move federal states for their vocational training (Socio-Economic Panel (SOEP), own calculations). Likewise, only 5% of trainees commute between federal states (Linked Employer-Employee Data of the IAB (LIAB), own calculations).

training degree.[3] Second, its institutionalized set-up allows for precise identification of trainees, training firms, and training periods in administrative data. Third, trainee wages are highly rigid (Muehlemann et al., 2022). Consequently, labor supply shocks are unlikely to engender wage responses (as I also document in the case of the missing trainee cohort), leading to below-equilibrium employment just as the definition of a labor shortage stipulates. Fourth, the low geographic mobility of trainees greatly enhances the sharpness of the negative trainee supply shock with respect to state boundaries, aiding identification.

I compare investments and technology adoption of firms in treated East German states undergoing the trainee shortage to investments and technology adoption of firms in control East German states not experiencing a trainee shortage in a difference-in-differences event study design. To ensure that no concomitant industry-specific shocks drive the results, I ensure comparability of treated and control firms by matching each treated firm to a comparable control firm operating in the same sector. I focus on training firms, defined as firms that employed trainees from the reformed school track prior to the reform. Non-training firms in non-exposed industries should not be directly impacted by the shock and serve in a falsification test. I use a large and representative firm panel survey containing information on firm investments and technological change combined with firm-level employment information from social security records.

I provide three key empirical findings. First, the education reform produces trainee shortages. The reform has a substantial negative effect on firms' employment of trainees from the reformed school track, i.e. trainees with a 12 or 13 years of schooling and a university entrance degree, henceforth "highly educated trainees". Highly educated trainees make up 16% of all trainees (Federal Statistical Office, Genesis-Online, 2022a), while the majority of trainees have 9 or 10 years of schooling, henceforth "low-educated trainees." Training wages do not increase. Firms also do not compensate missing highly educated trainees by hiring more low-educated trainees of workers with completed vocational training, commuting of trainees across states does not intensify, and internal training of incumbent workers is not expanded.

A second key finding is that trainee shortages cause reductions in investments: investments decrease sharply in training firms in treated states compared to training firms in control states in face of the trainee shortage. This implies that trainees are complementary to firm investments, While highly educated trainees represent 11% of a training firm's yearly hires and 3% of a training firm's workforce, investments per worker drop by €3,370 on average in affected years, corresponding to a decrease of almost 20%. The magnitude of the effect can be explained by firms refraining from *large* investment projects. This finding is in line with a literature emphasizing the lumpy nature of investments (e.g. Doms & Dunne, 1998; Cooper et al., 1999). Importantly, the reform-induced investment reductions are not recouped in subsequent years. This implies that even a temporary trainee supply shortage can lead to permanent reduction

[3] Based on the Sample of Integrated Labor Market Biographies (SIAB), own calculations.

in firms capital stock.

I confirm the link between the investment decline and the absence of trainees in several ways. First, non-training firms in non-training industries do not reduce their investments. Further, firms compensating the lack of trainees by hiring non-trainees also reduce investments, indicating that the investment response is specific to the shortage of trainees rather than a general labor shortage. Next, in an auxiliary identification strategy, I exploit firm-level exposure to the reform defined as pre-reform trainee employment. Confirming the complementary relationship between trainees and investments, firms more affected by the negative trainee supply shock decrease investments to a greater extent. This finding also ensures that the investment decline based on the event study design is not merely the result of firms' selection into hiring trainees despite the shortage.

A third key finding is that the induced investment decline – happening in a period of strong technological advancements, substantial investments in digital tools, software and computer-controlled machines, and changing skill requirements – is linked to reduced technology adoption. The negative trainee supply shock causes the technical condition of machinery to depreciate in treated training firms, and reduces investments in production technologies and information and communication technologies (ICT). Further, there is a substantial decrease in firm-level organizational change, which often accompanies technological shifts such as IT-driven workplace restructuring (Bresnahan et al., 2002).

To rationalize the effects of the negative trainee supply shock on firms' technology adoption, I present a stylized economic framework of endogenous technological change that incorporates capital adjustment costs of worker training. Training is necessary since new technologies are assumed to require new skills. As a key novel implication, technological change is endogenous to factors entering the capital adjustment costs. Opportunity costs of training young labor market entrants in terms of foregone production are lower for young labor market entrants because they are initially very unproductive, and concomitant productivity gains of training are large. Young labor market entrants hence complement technology adoption not due to their age but due to their comparative advantage in skill acquisition. When young labor market entrants are scarce, firms may not be able to adopt new technologies because adjustment costs of retraining incumbent workers are prohibitively high.

I provide empirical evidence in support of the mechanism via adjustment costs of worker training in new tech skills. The hypothesized mechanism implies that firms with higher shares of incumbent workers with outdated skills, and thus higher needs for trainees to acquire new skills, reduce technology adoption more when trainees are scarce. Based on Lipowski et al. (2024), who demonstrate that vocational training curricula are updated to include up-to-date technological skills, I provide empirical evidence for this implication. Also in line with the hypothesized mechanism, the investments drop is more pronounced in firms with higher trainee retention rates, i.e. firms employing trainees as an investment in skills for future production.

By being the first to show that shortages of young labor market entrants causally and significantly decrease firm technology investments, I contribute to three related literatures.

The most closely related strand of literature studies how technology invention and adoption responds endogenously to the relative abundance of production factors (e.g. Zeira, 1998; Acemoglu, 1998, 2002). Empirical papers, mainly relying on migration shocks, support this theory. For example, a decrease in the supply of *low*-skilled labor increases labor-saving patenting and fosters the adoption of labor-saving production technologies (Lewis, 2011; Hornbeck & Naidu, 2014; Clemens et al., 2018; Dechezleprêtre et al., 2019; Danzer et al., 2020; Andersson et al., 2022; San, 2023, also the other way around). In turn, increased supply of *high*-skilled labor intensifies the adoption of skill-complementing technologies (Beaudry et al., 2010; Carneiro et al., 2022). Endogenous technological change has also been shown to occur in response to demographic change: countries with lower population growth or shortages of middle-aged workers adopt more robots (Abeliansky & Prettner, 2017; Acemoglu & Restrepo, 2022). Above a certain tipping point, however, the lack of young workers reduces investments in information and communication technologies (Angelini, 2023). This paper's contribution to the literature on endogenous technological change is twofold. First, it incorporates capital adjustment costs consisting of worker training into such models, demonstrating that technology adoption is endogenous to factors entering capital adjustment costs, which can produce substantially different results. Second, it provides empirical evidence in a new, complementary setting, studying a negative supply shock of young natives, benefiting from a clear identification that is free from potentially confounding labor demand effects common to migration or fertility shocks, and moving to the firm level. The findings suggest that the effects of shortages of young workers on economic growth are likely more detrimental than previously surmised.

Second, I contribute to a literature on technology-specific human capital, the creation of new tasks, and their impacts on workers of different ages. The literature provides many examples of how new technologies require new skills and new tasks, without ruling out the replacement of existing tasks (e.g. Chari & Hopenhayn, 1991; Autor et al., 2003; Acemoglu & Restrepo, 2018; Autor et al., 2022). Such new skills have been linked to decreasing returns to experience, early retirement, and reduced hiring opportunities for older workers (e.g. Aubert et al., 2006; Ahituv & Zeira, 2011; Deming & Noray, 2020). Therefore, the literature on human capital investments predicts that adaptation to technological change takes place through the entry of young workers, rather than by upskilling incumbent workers (e.g. MacDonald & Weisbach, 2004; Hobijn et al., 2019; Cavounidis & Lang, 2020; Adão et al., 2020), leading to jobs exposed to technological change "getting old" (Autor & Dorn, 2009). Adão et al. (2020) show that this particularly applied to the ICT revolution which I study in this paper. My paper flips the logic around, manifesting that technology-specific skills induce adjustment costs of technology adoption consisting of worker training, and hinder the adoption of technologies when young labor market entrants are scarce.

Third, I contribute to a nascent literature on the consequences of labor shortages on firm outcomes. Existing studies establish a negative effect on firm capital, sales, and productivity

(D'Acunto et al., 2020; Le Barbanchon et al., 2023; Sauvagnat & Schivardi, 2023). I provide detailed evidence on one mechanism through which reduced labor supply affects firm outcomes, namely via firm technology adoption.

The remainder of the paper is structured as follows. The next section provides an overview of the vocational training system in Germany and the education reform under examination. Section 3 describes the data. I present the difference-in-differences event study approach in Section 4, followed by the empirical results regarding the reform's impact on trainee employment (Section 5) and firm technology investments (Section 6). Section 7 presents a stylized economic framework highlighting the mechanism via adjustment costs of worker training and provides empirical evidence for it. Section 8 concludes.

2 The German vocational training system and the education reform

2.1 The German vocational training system

Vocational training is a key component of both the German education system and labor market, with 60% of the working population possessing such training (Sample of Integrated Labour Market Biographies, own calculations). Vocational training often prepares for occupations which commonly require bachelor's or associate degrees in other countries like the US. It is commonly provided within the dual system, which combines on-the-job training at a firm (3-4 days per week) with vocational schooling provided by the state (1-2 days per week). A high share of trainees remain at their training company upon training completion.

Vocational training usually takes three years. Adolescents start vocational training after graduating from one of the following three high-school tracks: the basic track (*Hauptschule*, 9 years of schooling) which qualifies for vocational training in blue collar occupations; the intermediate track (*Realschule*, 10 years) which prepares for any vocational training, including training in white collar occupations; or the upper track (*Gymnasium*, 12 or 13 years) which is required for university studies. Approximately a third of the upper track school graduates chooses to undergo vocational training,[4] such that in 2000, 16% of trainees had a university entrance degree (*Abitur;* Federal Institute for Vocational Education and Training, 2009). Trainees from the upper school track often work in media, financial service, or retail occupations.

Trainees rarely move or commute to their workplace: only 2.2% lived in a different federal state before starting their vocational training (SOEP, own calculations).[5] Among those firms

[4]There were approximately 200,000 university entrants and 100,000 vocational training entrants with university qualification in 2000 (Federal Statistical Office, Genesis-Online, 2022c; Federal Institute for Vocational Education & Training, 2002). Similarly, Heine et al. (2005) report that 28% of upper track graduates from 1999 had enrolled in university studies six months after graduation, while 21% had started vocational training. 32% were in civil or military service, hence pursuing vocational training or higher education with one year delay.

[5]Likewise, only 6.0% of new trainees have lived in a different location before starting the vocational training.

covered in the subsequent analyses, the share of trainees commuting from a different federal state is also low with approximately 5%.

Regarding the central aspects of this paper, vocational training can be compared with on-the-job training in other countries: trainees are hired by their training company, receive a working contract for the duration of their vocational training and a training wage. Training wages are often subject to collective bargaining agreements and are low.[6] In contrast to on-the-job training in countries, state-provided vocational schooling transfers external knowledge to firms, and nationally binding training curricula ensure that the training content is not firm-specific and current.

A representative firm survey suggests that trainees play a key role in firms' acquisition of new tech skills: when asked about their vocational training, 44.5% of the firms agree that it ensures the constant supply of new skills, 46.5% agree that it improves the firms' innovative capacity, and 43% agree that it enhances the firms' adaptability to market and technological changes.[7] Similarly, Schultheiss & Backes-Gellner (2022) show that in Switzerland, a country with a vocational training system similar to Germany, changes in training curricula result in firms being closer to the technology frontier.

2.2 The reform

Prior to German reunification in 1990, upper track school graduates underwent 12 years of schooling in East Germany (Mecklenburg-Western Pomerania, Brandenburg, Saxony, Thuringia, Saxony-Anhalt, East Berlin) and 13 years in West Germany. After reunification, in an effort to align the the two education systems, Brandenburg switched to 13 years in 1994, while Saxony and Thuringia retained the 12-year system. With the graduation cohort of 2001, Saxony-Anhalt and Mecklenburg-Western Pomerania transitioned from 12 to 13 years. The education reform was decided in May 1996 in Mecklenburg-Western Pomerania and in January 1998 in Saxony-Anhalt.[8] By lengthening the years of schooling, the reform increased the level of education. More importantly, because the last cohort completing 12 years graduated in 2000 and the first cohort completing 13 years graduated in 2002, the adjustment resulted in a missing upper track school graduation cohort in spring 2001. While two thirds of the missing upper track school graduates would usually opt for university studies, one third would subsequently start vocational training. For identification, I exploit this labor supply in upper track vocational

These numbers are based on all German countries for the years 1984–2017. While the sample size shrinks considerably when restricting the data to East Germany and years around the reform in 2001, the numbers remain low (3.4% inter-state movers to East German states across all years, 4.3% inter-state movers to East German states 1995-2005.

[6]The average monthly gross compensation agreed by collective bargaining was €555 in 2000 (Federal Institute for Vocational Education & Training, 2022).

[7]BIBB-Cost-Benefit-Survey 2000, East German firms only, own calculations.

[8]For more information on the education reforms, see Kühn et al. (2013); Helbig & Nikolai (2015). Between 2007 and 2013, all German federal states adopted to the 12-year system, with Saxony-Anhalt making the change in 2007 and Mecklenburg-Western Pomerania in 2008. My study therefore concludes in 2005, making the 2001 reform considered permanent for the time frame examined.

trainees in 2001, while abstracting from the years of schooling. I assign Saxony-Anhalt and Mecklenburg-Western Pomerania as treated states and the other four states as control states.[9] I focus on upper track school graduates who subsequently start vocational training instead of university students/graduates because vocational trainees postpone their labor market entry less and move or commute less across federal states, thus endorsing the credibility of the identification strategy.

Mecklenburg-Western Pomerania, located in the northeast of Germany along the Baltic Sea, is a predominantly rural and sparsely populated federal state with approximately 1.6 million inhabitants as of 2020. Its economy is defined by small and medium-sized enterprises engaged in agriculture, maritime industries, mechanical engineering, and tourism. The second treated state, Saxony-Anhalt, situated in central Germany with a population of around 2.2 million in 2020, features a comparatively more urban environment. Bordering Western Germany, this state is characterized by economic sectors such as the chemical industry, mechanical engineering, and automotive supply.

Figure 1 shows how the missing school graduation cohort translates into reduced trainee employment. Panel A shows a sharp drop in the absolute number of upper track school graduates in 2001 – in Mecklenburg-Western Pomerania from 6,400 to 300, and in Saxony-Anhalt from 9,400 to 400 – while the figures remain relatively constant in the control states.

Based on the official training statistics, Panel B shows that the missing school graduation cohort translates into a missing cohort of newly concluded training contracts. Since at that time, males in Germany had to do military service of 10 months when reaching the age of 18, the decline prolongs to 2002. The number of training contracts with highly educated trainees dropped by approximately 60% in 2001 and 2002 in treated states. However, this number should be taken with caution, since contracts also declined in control states around this time, although to a smaller extent and not sharply timed.

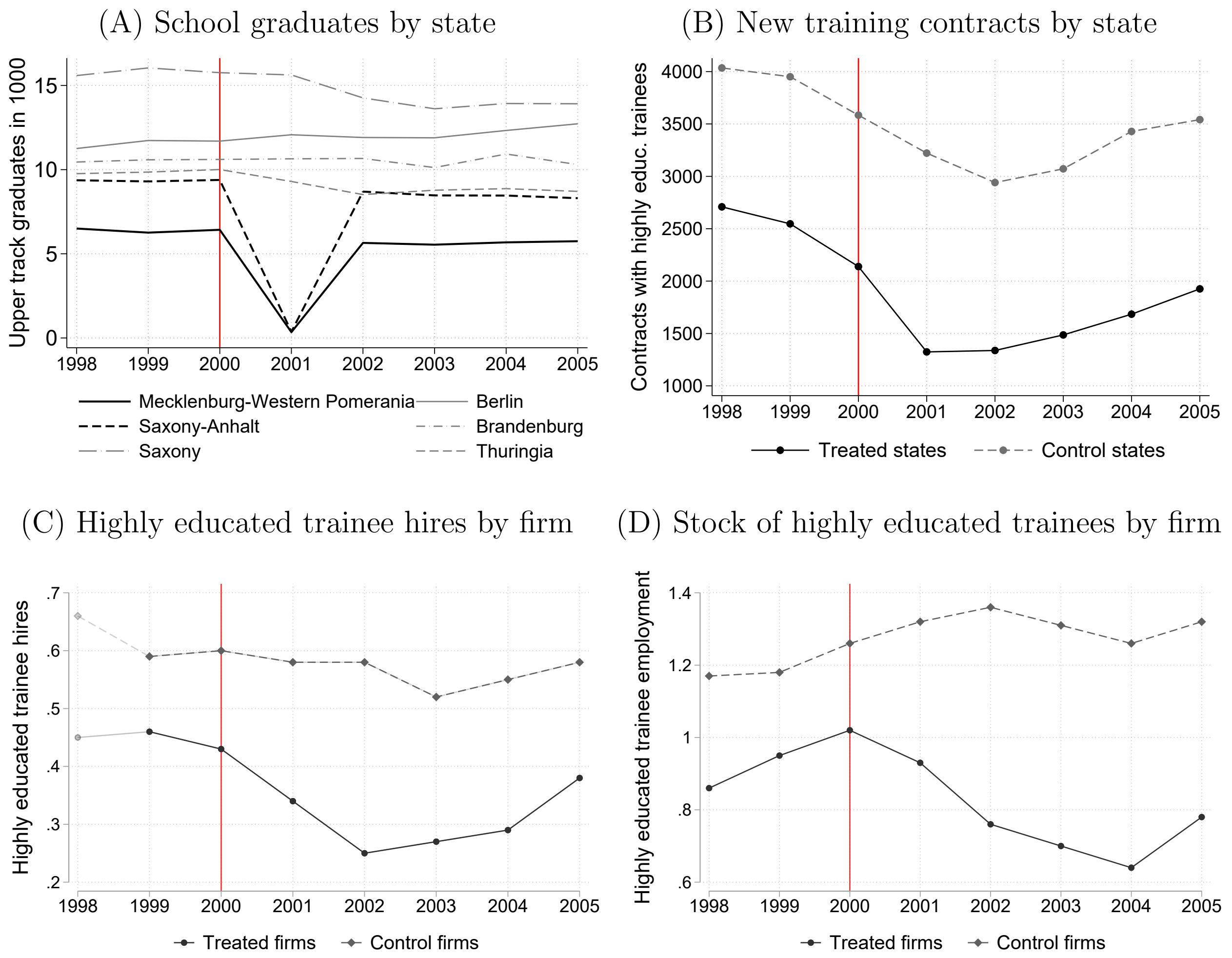
The missing trainees are also visible in the firm panel used for the subsequent investment analyses, see Panel C and D. The firm panel captures trainee employment with a lag of one year, because it is based on records as of June 30 each year, and most trainees start on August 01 each year. Hires of highly educated trainees decline by approximately 50% in 2002 in treated firms, corresponding to 0.16 trainees per firm.[10] Due to the usual training duration of three years, the stock of highly educated trainees is reduced by approximately one third in 2002–2004, corresponding to 0.34 trainees per firm.

The lack of highly educated trainees might also translate into fewer workers with completed vocational training several years later. However, this shock is largely mitigated over time, see Appendix Figure B1.1, Panel A. Likewise, there is no visible decrease in the number of workers

[9]The labor supply shock resulting from the missing upper track school graduation cohort in Brandenburg in 1994 has presumably dissipated until 2001.

[10]Hires in 1998 should be taken with caution since they are imputed based on observed employment. For firms entering the panel in 1998 it is hence impossible to determine whether an employee is a new hire or an incumbent worker.

Figure 1: The missing school graduation cohort



Notes: Panel A: Federal Ministry of Education & Research (2022). *Panel B:* Average number of new training contracts with graduates from the upper track in treated states (Mecklenburg-Western Pomerania and Saxony-Anhalt) and control states (Berlin, Brandenburg, Saxony, Thuringia). Including trainees of the dual system only. Federal Statistical Office, Genesis-Online (2022a). *Panel C and D:* LIAB, including trainees of the dual system only. Based on the balanced and imputed firm sample as described in Section 3. Hirings in 1998 should be taken with caution. Own calculations.

with tertiary education in affected states compared to control states, see Appendix Figure B1.1, Panel B, probably due to their high mobility across federal states and their tendency to postpone labor market entry.[11]

3 Firm panel data

Data sources. My analysis is based on the Linked-Employer-Employee-Data of the IAB[12] (LIAB), which combines the IAB Establishment Panel survey with administrative employment

[11] In the SOEP, 24.6% of all tertiary educated workers in East Germany have lived in a different federal state at the age of 17.

[12] IAB: Institute for Employment Research.

information of all employees at surveyed firms.[13] The IAB Establishment Panel is a large annual representative survey of establishments that includes information about investments, organizational change, sales, and internal training, among others. The Establishment Panel has existed in West Germany since 1993 and in East Germany since 1996. The number of surveyed establishments has risen from 4,000 in 1993 to 16,700 in 2020. Importantly, the survey is conducted at the workplace level, enabling the distinction between treated and untreated establishments based on their location.[14] I use the terms "firm" and "establishment" interchangeably for simplicity. Employment information is based on administrative records reported to the social security insurance. While employment information is reported as of June 30 each year, most vocational training programs start in fall, such that new trainees usually appear in the data with a lag of one year.

The data are well-suited for analyzing trainee shortages at the firm level because they provide a reliable distinction between trainees and workers with completed vocational training, in addition to wages and employment status. Also, information on schooling allows me to distinguish "highly educated" from "low-educated" trainees, i.e. trainees with a university entrance degree and those with a lower schooling degree, respectively. This is important since the education reform directly affects highly educated trainees only.

Data preparation. I restrict the dataset in four steps. First, I limit the data to firms in East Germany including Berlin, since the reform affects firms in East Germany which are not comparable with firms in West Germany. Second, I exclude firms in the health/education/social services sectors because vocational training in many related occupations is purely school based. The reform therefore does not affect firms' trainee employment in these sectors. Third, I limit the sample to firms with at least ten employees each year, as larger firms usually possess more accurate data and more consistent behavior over time. The results are robust to including smaller firms. Last, I constrain the sample to a balanced firm panel containing firms existing and with non-missing investments for the entire time period 1998–2005. A balanced panel has two main advantages over an unbalanced panel. First, it reduces compositional differences in the event study estimates that would likely violate the parallel trends assumption. Second, the firm-level matching procedure is only meaningful if treated and matched control firms are observed in the same years. As a disadvantage of a balanced panel, firms exiting the market or firms with missing values due to survey non-response are dropped. I find that firm exit is not affected by the reform, see Section 6. Conditioning on firm survival should hence not bias the estimates. With respect to survey non-response, I impute missing values by exploiting the panel dimension of the data. I proceed in two steps. First, I linearly interpolate

[13] I use the LIAB cross-sectional model which comprises employment spells that encompass June 30 of each year. The LIAB longitudinal model includes all spells but is unsuitable for this analysis because it is available for firms surveyed during the time period 2009–2016 only.

[14] The data does not allow to assign establishments to parent companies, precluding a within-company cross-establishment design.

missing values in up to two consecutive years if the firm has valid entries before and afterwards. This corresponds to imputing 2.1% of investment values but preserves an additional 13.0% of balanced firms. Second, I constantly extrapolate values at the start (1998, 1999) and at the end of the observation window (2004, 2005) for firms existing in these years as indicated in the social security records. At the cost of imputing 9.8% of investment values, this allows me to keep another 83.1% of balanced firms. Overall, by imputing 12% of investment values, the imputation procedure enables the inclusion of more than twice as many firms. For training firms, the share of imputed values is even lower. See Appendix A.1 for more details regarding the imputation procedure. The imputation procedure successfully recovers small firms with smaller investments which otherwise would have been lost due to the balancing requirement, enhancing the representativeness of the sample. I compute robustness checks which confirm the results in the non-imputed and/or unbalanced dataset.

Summary statistics. The final sample comprises 1,386 firms, of which 463 are treated and 923 are untreated. Table 1 shows summary statistics of the final dataset. In sum, all firms cover approximately 3.5% of the East German workforce in a year.[15] I observe on average 11,396 trainees per year, of which 1,558 (13.7%) are highly educated, corresponding to 1.12 highly educated trainees per firm, or 0.61% of a firm's workforce . Common occupations for highly educated trainees are media service occupations, retail occupations, insurance and financial service occupations, or technical drawer. Highly educated trainees are hence most common in the business service sector, but can also be found in the manufacturing or construction sector, see Appendix Figure A2.1.

Table 1: Summary statistics

	Mean	SD	Min	Max	Yearly sum
# workers	148	325	10	9,570	205,116
# trainees	8.22	27.76	0	846	11,396
# highly educated trainees	1.12	4.1	0	60	1,558
% highly educated trainees in total employment	.61	1.96	0	41.67	
No highly educated trainee	0.78	0.42	0	1	
No highly educated trainee 1998–2005	0.59	0.49	0	1	

Notes: SD: standard deviation. Yearly sum: Sum of workers across all firms.

Training versus non-training firms. In 78% of the firm-by-year observations, no highly educated trainee is employed, and 59% of the firms never employ a highly educated trainee over the entire time window 1998–2005, see Table 1. Since the reform affects highly educated trainee employment only, I focus on training firms, defined as firms with at least one highly educated trainee in 1998. This divides the sample into 293 training firms and 1,093 non-training

[15] The yearly average working population in East Germany from 1998 to 2005 was 59,406,800 according to Landesamt (2023).

Table 2: Pre-reform descriptives of training versus non-training firms

	Non-training firms N=1,093	Training firms N=293	Δ
# workers	110.97	354.32	-243.35^{***}
# trainees	4.62	22.37	-17.75^{***}
# highly educated trainees	.06	5.04	-4.98^{***}
% highly educated trainees in total employment	.06	2.45	-2.39^{***}
% highly educated trainee hires in total hires	0.63	10.76	-10.13^{***}
Inv. per worker (in €1,000)	14.61	18.80	-4.20^{***}
	Selected Industries		
Manufacturing	.33	.29	$.04^{**}$
Construction	.11	.06	$.05^{***}$
Business services	.11	.19	$-.08^{***}$
Public administration	.16	.25	$-.09^{***}$

Notes: Average values across 1998, 1999 and 2000 of training and non-training firms. Δ: Average in non-training firm - average in training firms. Selected indutries: those with a significant difference. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

firms. I base this classification on 1998 to minimize potential anticipation concerns. Less strict definitions of training firms, i.e. firms with at least one highly educated trainee in 1998 or 1999, or firms with at least on trainee in 1998 independent of the trainee's education show mitigated effects, as expected. Non-training firms in non-training industries are used in a falsification test. Table 2 shows summary statistics for training and non-training firms. In years prior to the reform, highly educated trainees made up 2.45% of a training firm's workforce and 10.76% of a training firm's hires. Training and non-training firms are fundamentally different. Compared to non-training firms, training firms are three times as large in employment, make larger investments per worker, and operate more often in the business service and public administration sector, and less often in manufacturing and construction.

Investments and technological change. Each year, firms in the Establishment Panel are asked whether they invested in four investment categories in the last year: (1) production facilities, plant and equipment, furnitures and fixtures, which I will term "production facilities", (2) communication technology, electronic data processing; "information and communication technologies (ICT)", (3) real estate and buildings; "real estate", and (4) means of transport, transportation systems; "transport". If a firm invested in at least one of these, the firm is surveyed on the total amount of annual capital investments. Accordingly, the investment volume is expected to contain investments in these four categories, while it is unlikely that non-tangible assets other than ICT are included. Finally, a firms is asked what share of total investment was attributed to the expansion of the establishment. In contrast, investments not targeted at the expansion of the establishment commonly include so-called start-up investments, replacement investments and rationalization investments, hence including investments related to technology adoption. Appendix Table A2.1 provides a detailed description of the underlying survey questions and variable construction.

To ensure that large firms with large investments do not drive the results, I focus on investments per worker, defined as total nominal investments divided by the initial number of workers in 1998.[16] The distribution of investments per worker is highly right-skewed, see Figure 2, Panel A, in line with the literature emphasizing the lumpy pattern of investments (e.g. Bessen et al., 2023). While 19% of the observations show no investment, the median investment per worker and year in the sample is €5,045 (€331,633 for total investments), the mean €17,641 (€2,679,418 for total investments), and the maximum investment per worker exceeds €300,000 (€15,000,000 for total investments). Investments per worker are highest in the energy/water/waster sector, followed by public administration and business services; and lowest in construction and hospitality, see Appendix Figure A2.2, Panel A. To curtail the influence of extremely large investments, I cap values in the top percentile of either total investments or investments per worker. Investments start with values as low as €6 per worker, which justifies treating the variable as continuous.On average, 35% of investments were attributed to firm expansion.

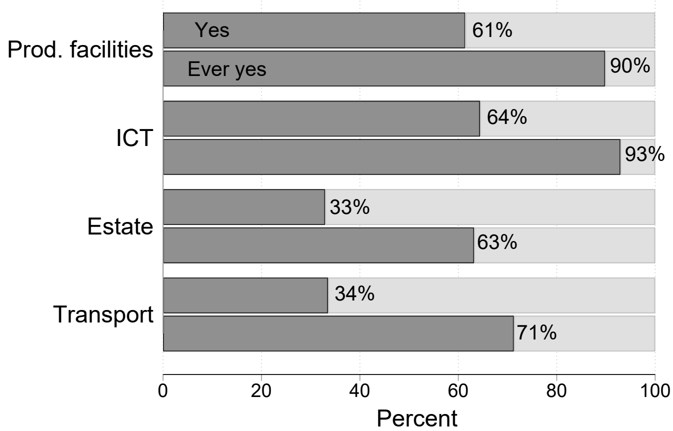
Figure 2, Panel B shows the share of firms investing in one of the four investment categories per year and at least once over the time window 1998–2005. More than 90% of the firms invest in production facilities or ICT at least once between 1998 and 2005. In 61% of the firm-by-year observations, I observe an investment in production facilities, and in 64% an investment in ICT. These shares do not vary strongly by industry, see Appendix Figure A2.2, Panel B. Investments in real estate and transport occur less frequently with one third of firm-year observations making an investment in either category. 3.5% of all firms never invest between 1998 and 2005. This share is lower among training firms (2%). In summary, capital investments occur regularly in the data.

Figure 2: Descriptives on firm investments

(A) Investments per worker

Share of obs. w/o investments	19.3%
5th percentile	€211
25th percentile	€1,391
50th percentile	€5,045
75th percentile	€17,661
95th percentile	€78,493
Mean	€17,641

(B) Share of observations by investment type



Notes: Panel (A): Observations at the firm-year level. For the corresponding table for total investments, see Appendix table A2.2. *Panel (B):* Yes: Firm-year observations with investment in a certain investment category. Ever yes: Firms invested in a certain investment category in any year between 1998–2005.

I use information on the firms' technical status of machinery and organizational change to

[16] Since inflation affects all firms equally, it will be absorbed by the year fixed effects.

directly measure technological change. Out of all the firm-year observations, 0.4% rate the technical status of machinery as the lowest category 1, which corresponds to 'completely out-of-date.' 3% assess it as category 2, 30% as category 3, 51% as category 4, and 16% as the highest category, labeled 'state-of-the-art.' There is variation in technical status within firms over time: In 30% of the observations, firms' technical status changes from one year to the next.

Firms report whether they implemented organizational changes, which often complement technological change. I follow Battisti et al. (2023) and define organizational change on a scale from 0 to 4 by adding up four binary indicators. These indicators are 1) restructuring of departments or areas of activities, 2) downward shifting of responsibilities and decisions, 3) introduction of team work/working groups with their own responsibilities, and 4) introduction of units/departments carrying out their own cost and result calculations. More than half of the firm-by-year observations report none of the four changes, 22% report one change, 12% two changes, 5% three changes, and 1% four changes. In 43% of the cases, firms' technical status changes from one year to the next.

Table 3: Technological change and investment types

	Δ Technical status		Organizational change	
	(1)	(2)	(3)	(4)
Production facilities	0.02	0.02	0.12***	0.09***
	(0.01)	(0.01)	(0.03)	(0.03)
ICT	0.03**	0.03**	0.29***	0.20***
	(0.01)	(0.01)	(0.03)	(0.03)
Real estate	0.03***	0.03***	0.05	0.02
	(0.01)	(0.01)	(0.04)	(0.04)
Transport	−0.01	−0.01	0.02	0.04
	(0.01)	(0.01)	(0.03)	(0.03)
Year FE	✓	✓	✓	✓
Base controls		✓		✓
Observations	9699	9699	5053	5053

Notes: Base controls include industry fixed effects, firm employment size categories and federal state dummies. Investment type lagged by one year. Organizational change is observed in the years 1998, 2000, 2001, 2004 and 2007 only. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

I next analyze which investment categories embody technological change. To do so, I regress changes in the technical state of a firms machinery, i.e. firm-level technological change, and organizational change on each of the investment categories, controlling for year fixed effects. The results are shown in Table 3. In columns (2) and (4), I additionally control for industry, firm size, and state. Investments in ICT and real estate are significantly positively related to changes in a firm's technical status, while investments in production facilities and ICT are positively associated with organizational change. Investments in transport are not correlated with any of the two measures. I therefore interpret investments in production facilities and ICT as embodying technological changes; investments in real estate as complement to technological change, and investment in transport as a placebo outcome unrelated to technological change.

4 Event study approach

The identification strategy exploits exogenous variation in the supply of upper track school graduates across states and years in a difference-in-differences event study design. I compare treated and control firms before and after the reform by estimating the following specification:

$$Y_{jbt} = \sum_{t=1998}^{1999} \alpha_t(\text{Treated}_{b(j)} \times \text{Year}_t) + \sum_{t=2001}^{2005} \beta_t(\text{Treated}_{b(j)} \times \text{Year}_t) + \psi_t + \phi_{b(j)} + \epsilon_{jt} \quad (1)$$

where Y is one of several outcomes such as investments, j denotes the firm, b the federal state, and t the calendar year. Treated is a binary variable with Treated $= 1$ if the firm is located in a state undergoing the reform and zero otherwise. ψ_t captures calendar-year fixed effects common to both treated and control states. State fixed effects $\phi_{b(j)}$ capture time-constant level differences between federal states. The results are robust to including firm fixed effects. The vector β_t includes the coefficients of interest, namely the differential investment behavior of firms in treated states compared to firms in control states following the reform in 2001. I stop in 2005 because of a different education reform affecting trainee supply from 2007/2008 onwards. Note that treatment is not staggered, precluding potential biases common to two-way fixed effects estimators in a staggered setting (e.g. Goodman-Bacon, 2021).

I estimate equation (1) for training firms and non-training firms separately. A firm is classified as a training firm if it employed at least one highly educated trainee in 1998, and as non-training firm otherwise. The reform has a direct impact on training firms, while non-training firms are unaffected, except for spill-over effects. I therefore use non-training firms as a falsification test and expect much smaller estimates.

The identification of the causal effect in the event study relies on three main assumptions.

Assumption 1 - Parallel trends. First, I assume that firm outcomes in treated states in absence of the reform would have evolved in parallel to those in control states. A common approach to evaluate the credibility of this assumption is to check for parallel trends prior to the shock, as I do in Sections 5 and 6. To ensure that no change in firm composition violates this assumption, I restrict the data to a balanced panel with non-missing investments for the entire time window 1998 to 2005.[17]

[17] Firm exit might be related to trainee shortages caused by the education reform. Conditioning on firm survival might therefore result in a lower bound of the estimated reform-induced investment drop since exiting firms are likely those which would have invested little or not at all, had they survived.

One might be concerned that treated training firms differ from control training firms in aspects which expose treated training firms to different potential confounders than control training firms. Also, the treatment effect might be heterogeneous across firm characteristics, which could bias the coefficients of interest if certain firm characteristics are more prevalent among treated firms than among control firms. Indeed, treated training firms operate less often in manufacturing than control training firms. In terms of other conceivably relevant pre-reform firm characteristics, such as the share of highly educated trainees or investments, treated training firms and control training firms differ remarkably little, see Appendix Tables B1.1 and B1.2. To ensure that treated and control training firms are comparable, I match treated and control firms based on their pre-treatment characteristics. The matching procedure consists of two steps. In a first step, I match firms within training and non-training firms and within nine broad industry groups. By matching within industries, the estimated reform effects are devoid of distorting industry-specific shocks or heterogeneity in treatment effects by industries. In a second step, I perform Mahalanobis distance matching with replacement. This metric minimizes the standardized Euclidean distance of the matching variables between treated and control firms, while taking into account the correlation between the matching variables. The matching variables include pre-treatment log overall employment, pre-treatment relative employment of highly educated trainees and pre-treatment investments per worker. I directly match on investments in all pre-treatment periods since investments cannot be well approximated by other covariates due to their lumpiness (Bessen et al., 2020). Because restricting on firms with no pre-trends is problematic in case there are pre-trends (Roth, 2022), I report results for the sample of all firms and for the sample of matched firms throughout the paper. Convincingly, the results are similar for both samples. The results are robust to the matching specification.

The matching procedure does not provide remedy if external factors evolve differently in treated and control states. I therefore check that population growth and the unemployment rate exhibit comparable patterns across states, see Appendix Figure B1.2.[18] Moreover, one might be concerned that the introduction of the euro in 2002, the German Hartz reforms over 2003–2005, the bust of the dot-com bubble in 2000, or China's accession to the World Trade Organization in 2001/2002 might confound the reform effect. However, these shocks likely affected treated East German states and control East German states similarly, especially within industries. In addition, it is unclear why any other shock would affect firm outcomes differently based on the share of highly educated trainees at a firm.

[18]While there was a notable outflow of workers out of East Germany following the fall of the iron curtain in 1989, this affected treated and control states similarly. To avoid unintended distortions because population growth might react to the reform, I look at the population of 14-years old four years earlier. If any, Berlin and Brandenburg show slightly different patterns. Robustness checks excluding these two states provide very similar results. Regarding the unemployment rate, Saxony shows a slightly distinct trend. I therefore exclude Saxony in a robustness check which does not affect the results.

Assumption 2 - No Anticipation. The second identifying assumption is that firms did not change their behavior prior to the reform. Since the reforms were decided in 1996 and 1998, firms had the opportunity to adjust their employment and investments prior to 2001. However, the event study estimates show little evidence for this.[19] Students might have also anticipated the reform. When the reform was decided, students of the missing graduation cohort were in grade 7 in Mecklenburg-Western Pomerania and in grade 9 in Saxony-Anhalt. Since the choice of school track was due after grade 6 in East Germany, it was not impacted by the reform. If school graduates might delay or accelerate the start of their vocational trainings, this would bias the estimates towards zero.

Assumption 3 - No spill-overs/SUTVA. Third, I assume that control states are not affected by the reform, and treated states are not affected by the absence of the reform in control states. This assumption is violated if trainees move or commute across federal states. The data allows to identify cross-state commuting. Trainees rarely commute (5% in 1999 to 2001) compared to workers with a university degree (9%), and this share does not change in response to the reform, see Section 5. To investigate whether school graduates move for their apprenticeship, I turn to the Socio-Economic Panel (SOEP) which tracks individuals from childhood onward. The cross-state trainee moving rate is extremely low with 2.2%. Also, there is no instance of a highly educated trainee relocating to one of the treated federal states in the post-reform years 2001, 2002 or 2003 in the data. However, even if trainees moved or commuted from control states to treated states in response to the reform, this would bias the estimates of towards zero.

Interpreting the reform as supply shock of trainees. Beyond identifying the causal impact of the reform itself, I aim to attribute the effects on firm investments to the decrease in trainee supply. This requires that no other aspect of the reform affects investments.

A potential confounding aspect of the reform is that highly educated trainees before the reform are different from highly educated trainees after the reform due to the increased years of schooling. However, the investment drop is only temporary, indicating that it is unlikely to be caused by permanent secondary aspects. Also, higher levels of education would, if any, likely induce more investments, and therefore provide a lower bound of the effect.

Potential concomitant demand shocks common to labor supply changes due to migration should also be unproblematic since the overall population size remains unchanged, and per capita spending likely adjusts only marginally since trainees earn very low wages.

Another concomitant factor is the potential substitution of missing trainees with workers of a different observed or unobserved type.[20] However, I do not interpret a substitution of missing

[19] With respect to investments, firms' anticipation of the reform might have actually afforded them time to recalibrate their plans, enabling investment reductions when the shock eventually materialized.

[20] Trainees starting in 2001 are likely negatively selected in terms of unobserved characteristics: they come from an unaffected state, from a previous graduation cohort, or forego university studies to start vocational

trainees as a source of bias but as a mechanism via which the effect unfolds. Also, given that any replacement was far from complete, I am confident that a large fraction of the estimated effect on investments is due to a drop in the supply of trainees, rather than a change in worker composition.

Trainee distribution across firms. Even if the estimated parameter of interest $\hat{\beta}_t$ is unbiased, it is subject to the distribution of trainees across treated firms. In particular, $\hat{\beta}_t$ is small if trainees are primarily missing in firms that would not have invested in absence of the shock, for example never-investors, even though only 2% of the training firms never invest over the observed time period. $\hat{\beta}_t$ is large if trainees are primarily missing among firms that would have invested in absence of the shock. In order to identify the effect independent of the realized distribution of trainees across firms, I propose a complementary identification strategy in Appendix C: I predict the distribution of trainees across firms based on a Bartik style instrument consisting of firms' pre-reform use of trainees and the state-level shift in trainee employment.

Inference. In general, standard errors should be clustered at the level of treatment assignment to account for cluster-level shocks (e.g. Abadie et al., 2023). In this setting, the number of clusters, i.e. federal states, is small. For valid inference with a small number of clusters, I follow three approaches. First, as suggested by Roth et al. (2023), I assume that the cluster-specific shocks are small compared to the idiosyncratic error terms at the firm level, potentially resulting in a small violation of parallel trends. The remaining uncertainty comes from the sampling of firms within clusters only. I hence cluster standard errors at the firm level. Second, I report confidence intervals based wild t-bootstraps clustered at the state level as suggested by Cameron et al. (2008). Third, I perform Fisher randomization (permutation) tests, computing the t-statistic of the treatment effect for the actual treatment assignment and for all permuted treatment assignments across federal states.

5 Bite of the reform

Figure 3 displays the results of the difference-in-differences event study model outlined in equation (1) regarding the effect of the reform on trainee hires (Panel A) and trainee employment (Panel B). The left panel shows the result based on the sample of all firms, the right panel based on the sample of matched firms. Panel A shows a clear drop in hires of highly educated trainee in treated training firms compared to control training firms in 2002. This is precisely the year the majority of the upper track school graduates of 2001 would have appeared in the data as new hires.[21] With 0.64 fewer hires in 2002 in the sample of all firms (-1.11 in the

training. Individuals with better unobserved characteristics likely do not need to follow any of these three strategies.

[21]Note that vocational training usually starts on August 1st each year, while firm employment is recorded as of June 30th each year, leading to a year delayed appearance of the missing school graduates in the data.

sample of matched firms), this corresponds to a pronounced drop of approximately 30% (50%). Hires in treated training firms remain slightly below hires in control training firms in 2003 and 2004, likely due to postponed entry related to military service. The gap has closed by 2005. Pre-trends in hires should be interpreted with caution due to the challenging identification of hires in the dataset in 1998.[22]

Panel B focuses on the stock of highly educated trainees, which experiences a longer-term decline since vocational training usually takes three years. In 2002, 2003 and 2004, approximately 1.5 fewer highly educated trainees work in treated training firms compared to control training firms. With an average of 4.9 highly educated trainees per training firm in 1998, this corresponds to a drop by one third. Considering the typical training duration of three years, this aligns with the absence of one year's worth of upper track school graduates. Consistent with the timeline of the shock, the employment gap starts to shrink in 2005.[23] Firms' highly educated trainee employment evolves in parallel in control and treated states in the years 1998 to 2000, likely because there are no suitably qualified trainees available to employ in anticipation of the reform. Trainee employment already starts to drop in 2001, potentially due to some trainees already being employed at their training firms on June 30 before the official training start on August 1st. Convincingly, the estimates are comparable across the sample of all training firms and the sample of matched training firms.

Wage and worker substitution effects. The data allows to study firm adaptation behavior, such as wage changes or the substitution of highly educated trainees with other workers. To investigate such accompanying effects, I employ the corresponding difference-in-differences specification, distinguishing the pre-treatment period 1998–2000 and the post-treatment period 2002–2004. Results are given in Table 4.

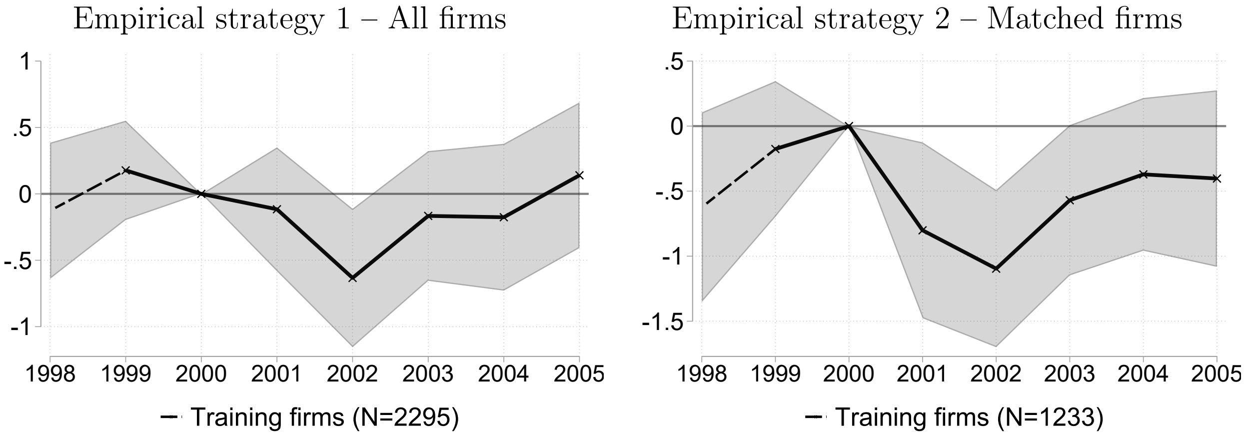
In contrast to what standard economic theory would predict, there is no evidence of an increase in wages of highly educated trainees in response to the negative supply shock (column 1). This result is not caused by firm selection into training highly educated trainees in the years of reduced trainee supply, as the specification with firm fixed effects reveals (column 2). This finding is in line with Muehlemann et al. (2022) and may be due to several reasons. First, firms likely shy away from increasing wages in response to a temporary supply shock because downward rigid wages will impede a subsequent wage decline as soon as the supply shock fades out. Second, supply of school graduates is by construction very inelastic. In order to attract non-school graduates, trainee wages would need to increase by a very large amount given their very low initial levels. Third, trainee supply reacts to expected wages post training graduation

[22]Since firms do not report new hires themselves, I impute hires based on observed employment. For firms entering the panel in 1998 it is hence impossible to determine whether an employee is a new hire or an incumbent worker.

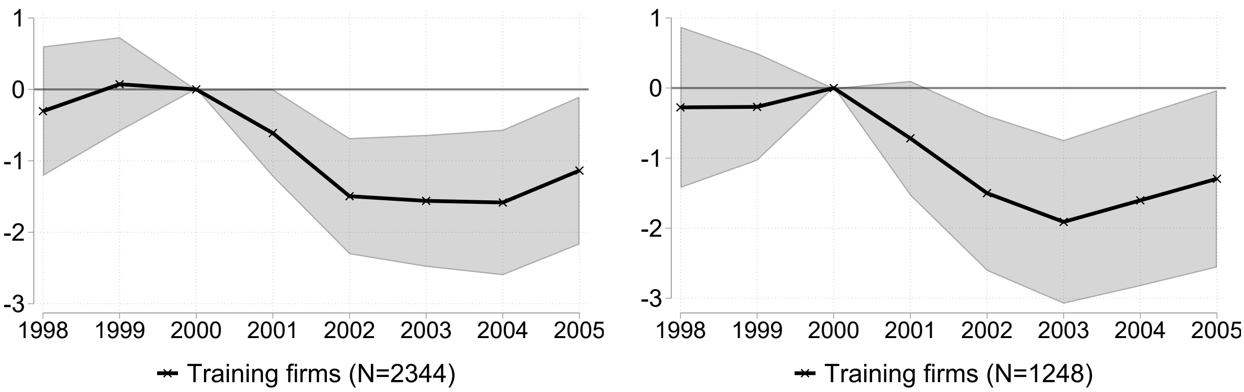
[23]I stop in 2005 to avoid confusion with a positive shock to trainee supply in 2007 and 2008, when Saxony-Anhalt and Mecklenburg-Western Pomerania switched back to the 12-year school system. This reform was unexpected at the time of the reform studied in this paper.

Figure 3: Effects on trainee employment

(A) Highly educated trainee hires



(B) Stock of highly educated trainees



Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Standard errors clustered at the firm level. Hirings in 1998 should be taken with caution. For the corresponding graph with confidence intervals based on cluster wild t-bootstraps, see Figure B2.1.

(Neuber-Pohl et al., 2023). Even a large percentage increase in training wages is negligible compared to unchanged expected wages post training graduation, given the comparatively short time of training and low initial training wages.

I now turn to potential worker substitution effects. Firms do not compensate for their missing highly educated trainees by hiring more low-educated trainees (column 3). In consequence, overall trainee hires also drop. The low substitutability between low- and highly educated trainees, also in line with Muehlemann et al. (2022), is likely related to distinct skill sets, the specialization in different occupations, and the stable demand for low-educated trainees.

Next, I study hires of highly educated trainees that commute from a different federal state (column 4). The coefficient of interest captures potentially increased commuting into treated states plus potentially reduced commuting into control states. There is no evidence of increased cross-state commuting of highly educated trainees following the shock, supporting the SUTVA assumption of no spill-overs across state borders.

Firms may also increase retraining of incumbent workers to overcome skill shortages. In

Table 4: DiD Results – Wage and worker substitution effects

	Wage effects		Substitution effects			
	Log wages highly educ. (1)	trainees (2)	# low-educ. trainee hires (3)	# highly educ. commuting trainee hires (4)	Internal retraining (5)	Log VT employment (6)
	Empirical strategy 1 – All training firms					
Treated × Post	−0.03	−0.04	−0.42	0.03	−0.09*	−0.13**
	(0.04)	(0.03)	(0.84)	(0.05)	(0.05)	(0.05)
	[-0.21;0.11]		[-0.75;2.93]	[-0.21;0.13]	[-0.35;0.07]	[-0.25;-0.08]
N	1758	1758	2295	2018	2227	2344
Firm FE		X				
Init. outcome	3.00	3.00	6.33	0.04	0.42	4.87
	Empirical strategy 2 – Matched training firms					
Treated × Post	0.01	−0.01	−0.04	0.04	−0.07	−0.07
	(0.05)	(0.04)	(1.11)	(0.06)	(0.07)	(0.07)
	[-0.16;0.14]		[-2.44;3.02]	[-0.15;0.27]	[-0.40;0.08]	[-0.17;0.07]
N	908	908	1233	1082	1190	1248
Firm FE		X				
Init. outcome	3.00	3.00	6.00	0.03	0.43	4.82

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Post: 2002–2004. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 90% confidence bands based on cluster wild t-bootstraps in square brackets. Init. outcome: Average outcome of treated firms in 1998. Column 4: a commuter is defined as a person living and working in two distinct federal states. This variable is available from 1999 onward. Column 5: Internal retraining is the sum of retraining incidences at the firm-year level. VT: completed vocational training. For the full set of results, see Appendix Table B2.1. For further outcomes see Appendix Table B2.2.

contrast, I observe a decline in internal training measures in treated training firms by approximately one third of the initial value (column 5). This finding might be related to foregone technology adoption and foregone organizational change, as I show below. Column 6 shows that employment of workers with completed vocational training does not increase in response to the trainee shortage, indicating that already trained workers are no suitable substitutes for trainees.[24]

To sum up, the reform leads to a sharp decline in employment of highly educated trainees that is not accompanied by higher trainee wages, and not compensated by low-educated trainees, increased commuting, retraining of incumbent workers, or increased employment of workers with already completed vocational training.

[24]Employment of already trained workers even decreases in training firms in treated states compared to training firms in control states. Disentangling hires and separations, I find that this joint effect consists of both increased hires and increased separations, see Appendix Table B2.2, potentially relating to an overall employment decrease, as discussed below.

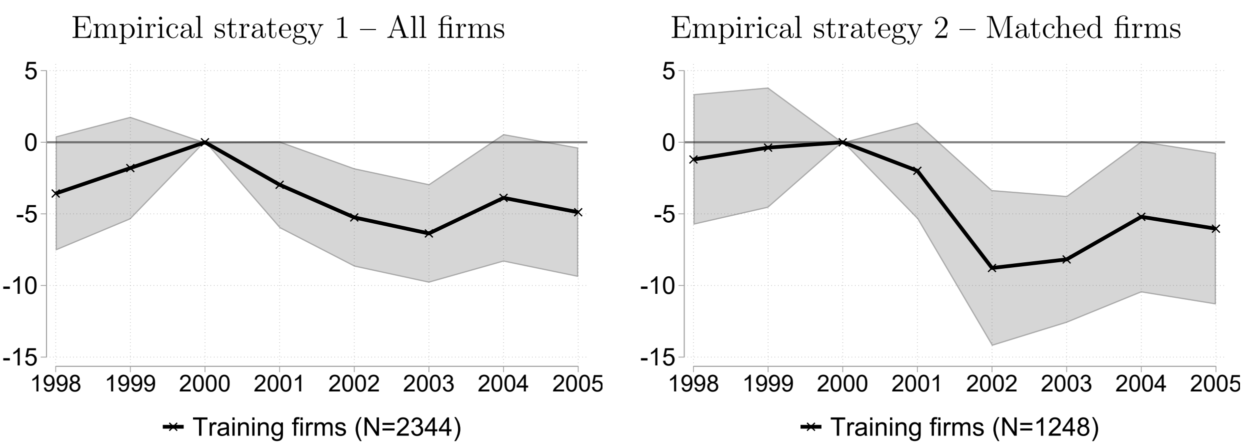
6 Effects on firm technology investments

6.1 Overall effect on investments

I now turn to the impact of the negative trainee supply shock on firm investments. Figure 4 shows a large and statistically significant decline in investments per worker following the reform in treated training firms compared to control training firms; the key finding of this paper. The strongest investment decline is observed in the years 2002 and 2003, with €5,250 and €6,370 less per worker in the sample of all firms. This corresponds to a drop of 30% and 37% of the average 1998 value. Within the sample of matched firms, the investment drop is even larger.

There are no statistically significant pre-trends for the sample of all firms.[25] No anticipation in investments is consistent with the idea sketched out in Section 7 that new technologies arrive constantly and firms are unable to adopt them once trainees are missing. The negative effect diminishes after 2003, corroborating its relation to the temporary drop in trainee supply. While it seems that investments do not fully return to their initial level by 2005, this finding is not statistically significantly different from zero and not robust across specifications. Importantly, the reform-induced foregone investments are not recouped at least until 2005. This implies that even a temporary trainee supply shortage leads to a permanent reduction in the capital *stock* (assuming no immediate capital depreciation).

Figure 4: Effect on investments per worker in €1,000



Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Standard errors clustered at the firm level. Outcome: investments divided by total employment in 1998 in €1,000. For the corresponding difference-in-differences estimate see Table 5. For the corresponding graph with cluster wild t-bootstrap confidence intervals, see Figure B2.2.

I employ alternative specifications of the outcome variable in corresponding difference-in-

[25] The sample of all training firms exhibits a slight investment increase before the reform. If firms anticipated the trainee shortage, they likely decreased expansion investments in anticipation of an employment decline, but increased technology investments in anticipation of the absence of easily trainable workers. Empirically, however, the investment increase prior to the reform is caused by increased expansion investments. This indicates that firms neither anticipated the employment decrease nor the incapability to adopt new technologies following the reform. Matching on pre-trends, i.e. on pre-reform investments per workers, is therefore a valid approach to remove these pre-trends.

differences regressions. The results are given in Table 5. First, I ensure that the effect is not driven by changes in the denominator, i.e changes in employment. When dividing investments by the current number of workers instead of the initial number of workers (column 2), I find a negative but smaller effect on investments per worker, indicating that the investment decline is partly but not exclusively driven by a concomitant employment decrease.

Next, I account for the highly right-skewed distribution of investments in combination with frequently observed zeros. Since a simple log-transformation has been acknowledged to be problematic (e.g. Chen & Roth, 2023), I instead employ several alternative transformations suggested by Chen & Roth (2023). In particular, I separate the effect on the extensive margin from the effect on the intensive margin (columns 3 and 4). This allows to log-transform investments for strictly positive values. Investments decline both at the extensive and at the log-transformed intensive margin, by 3–7% of their initial values. In order to combine both margins, I manually assign an importance to the extensive margin. In particular, I define a change from zero to any strictly positive investment to be as important as an investment increase by 1% (column 5). This combined measure indicates an investment drop of 8–10% of its initial value. Compared with other estimates in the literature, the estimated decrease corresponds approximately to the decline the literature would predict if capital costs permanently increased by 9-15% (Zwick & Mahon, 2017; Lerche, 2019; Liu & Mao, 2019).[26]

This finding suggests that trainees are complementary to investments. Interestingly, the investment decline is not only found for firms operating in business services and public administration, where one might expect labor to be complementary to technology, but also in the manufacturing sector, see Appendix Table B2.4, column 2 and 3.

Effect size. Since investments are divided by firm employment, the investment decline goes beyond a potential "mechanical" effect of reducing capital in proportion to trainee employment.[27] In fact, the average investment decline might seem large given that highly educated trainees present on average 10.7% of a training firm's yearly hires, and 2.5% of a training firm's workforce. To understand the magnitude of the estimate better, I plot the distribution of the underlying firm-level matched difference-in-differences estimates in Figure 5. The distribution of the treatment effect on investment is highly right-skewed. The *average* investment drop

[26] Zwick & Mahon (2017) study bonus depreciation in the US between 2001 and 2010, finding that a 1% reduction in investment costs increases investments by 3.69 log points. Lerche (2019) estimates an increase in investments by 2.43 log points in response to a 1% reduction in investment costs in the setting of investment tax credits in East Germany in 1999 among manufacturing firms. Liu & Mao (2019) find a value of 2.26 in China.

[27] In addition, a representative firm survey, the BIBB-Cost-Benefit-Survey 2000, suggests that the mechanic costs are much smaller than the estimated effect: East German firms surveyed in 2000 spent €487 on average per year and trainee on equipment and material (Beicht et al., 2004). With a reform-induced reduction in the number of trainees by 1.50 in 2002 and an average size of training firms of 354 workers this would imply a mechanic reduction of €2.06 per worker. In addition to these €487 on equipment and material costs, East German firms in 2000 reported €1530 of "other costs" per trainee per year, including costs for teaching material, fees, and training administration. If a firm interpreted all these costs as capital investments, the total mechanic reduction in investments would still be as small as €8.55 per worker.

Table 5: DiD Results – Investment effects

	Investments per worker		Intensive vs. extensive margin			
	per init. # of workers (1)	per current # of workers (2)	Any inv. (0/1) (3)	Log(Inv.) (4)	Combined (5)	Large inv. (1/0) (6)
			Empirical strategy 1 – All training firms			
Treated × Post	-3.37^{*}	-1.05	-0.05	-0.33^{**}	-0.65^{***}	-0.07
	(1.79)	(2.39)	(0.03)	(0.16)	(0.24)	(0.04)
	[-7.18;-0.13]	[-6.28;-0.94]	[-0.19;-0.02]	[-0.52;0.21]	[-1.49;-0.17]	[-0.16;0.08]
% of init. outcome	-19%	-6%	-6%	-4%	-10%	-13%
N	2344	2344	2344	2069	2344	2069
Init. outcome	17.43	17.43	0.89	7.45	6.62	0.54
			Empirical strategy 2 – Matched training firms			
Treated × Post	-6.86^{***}	-6.17^{*}	-0.03	-0.50^{**}	-0.55	-0.12^{**}
	(2.29)	(3.46)	(0.05)	(0.21)	(0.37)	(0.05)
	[-10.70;-1.72]	[-20.82;-0.88]	[-0.19;0.10]	[-0.74;-0.15]	[-1.54;0.49]	[-0.21;-0.01]
% of init. outcome	-39%	-35%	-3%	-7%	-8%	-22%
N	1248	1248	1248	1102	1248	1102
Init. outcome	17.68	17.68	0.88	7.41	6.55	0.55

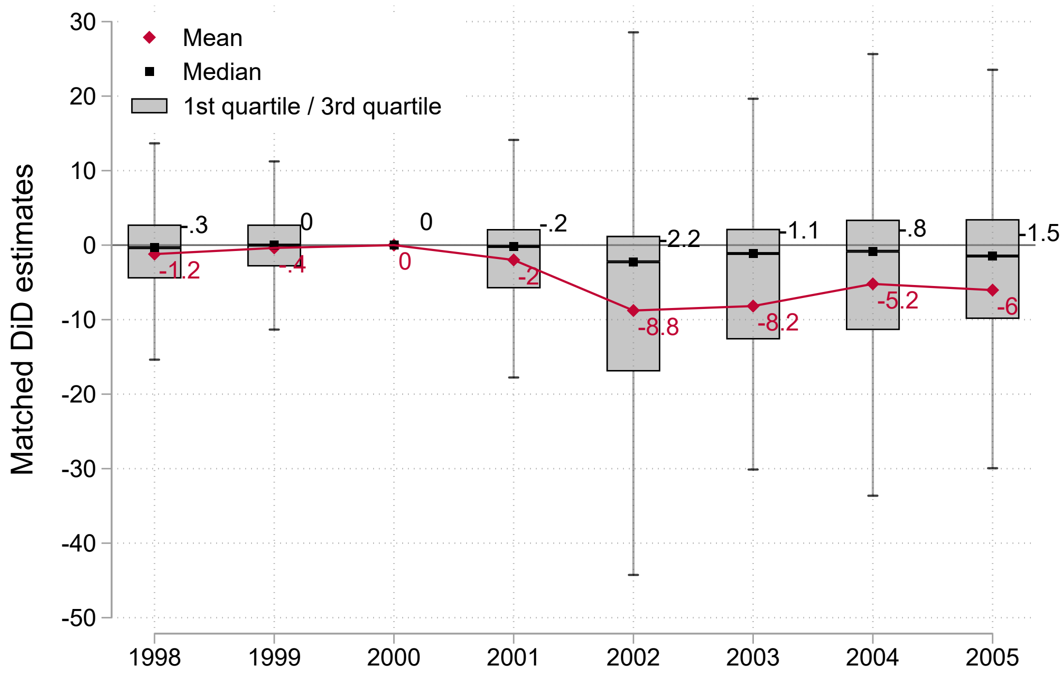
Notes: Reference group: Treated × Pre. Pre: 1998–2000. Post: 2002–2004. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 90% confidence bands based on cluster wild t-bootstraps in square brackets. Init. outcome: Average outcome of treated firms in 1998. For the full set of results, see Appendix Table B2.3. For further outcomes see Appendix Table B2.4.

in 2002 – the statistic reported throughout the paper – is four times larger (€-8,800; 50% of the 1998 value) than the *median* investment drop (€-2,200; 12% of the 1998 value).[28] The corresponding figure for total investments (Appendix Figure B2.4) shows a similar pattern. This finding implies that the investment drop is driven by the absence of large investments. This is in line with the literature highlighting the lumpy nature of investments (e.g. Cooper et al., 1999; Bessen et al., 2020), highlighting that firms often do not invest at all but if they do, investments are large. Applied to the present setting, some firms were not planning to invest even in the counterfactual scenario, so they do not reduce investments in response to the negative trainee supply shock. Other firms were planning to make a large, lumpy investment which they forego due to the reform.

To explicitly analyze the effect on large investments, I run a difference-in-differences regression among observations with strictly positive investments using a binary outcome taking the value one for investments in the upper tercile of the investment per worker distribution (>€10,000), and zero otherwise, Table 5, column 6. Treated training firms are 7–12 percentage points (13–22%) less likely to make large investments than control training firms when trainees

[28] Note that outliers in the uppest investment percentile are excluded. Also, the result is driven by a considerable number of missing large investments, suggesting that these are not merely outliers.

Figure 5: Distribution of matched DiD estimates – Investments per worker in €1,000



Notes: Distribution of the matched firm-level difference-in-differences estimates. Outcome: Investments per worker in €1,000. *Red:* Average. *Black:* Median. *Box:* 25th and 75th percentile. *Adjacent values:* 25th percentile - 1.5*the interquartile range (75th percentile + 1.5*the interquartile range). For the equivalent graph for total investments, see Appendix Figure B2.4.

are scarce. The effect is also negative but smaller when focusing on investments per worker in the upper decile (>€51,200), see Appendix Table B2.4, column 5, and when defining large investments within industries, the effect remains, see Appendix Table B2.4, column 6.

Falsification test among non-training firms. Next, I turn to the sample of non-training firms. Non-training firms, defined as firms with no highly educated trainee in 1998, should not be directly affected by the reform but might be indirectly affected via spill-over effects. Confirming this hypothesis, the average investment drop among non-training firms is less than half of the average drop among training firms, see Table 6, columns 1 and 2. See Appendix Figure B2.5 for the event study results. The investment drop among training firms is, however, not exactly zero. When using more restrictive definitions of non-training firms, i.e. firms never employing a highly educated trainee between 1998 and 2000, or firms never employing a trainee of any school education between 1998 and 2000, the small but negative effect on investments remains, see columns 4 and 6. This negative effect among non-training firms might stem from industry spill-overs, such as product market competition, knowledge spill-overs, inter-industry poaching of workers with completed vocational training, or from remaining mistakes of training firms as non-training firms. To test for industry spill-overs and to check that non-training firms in non-training industries do not decrease investments, I perform a difference-in-differences regression including the triple interaction term between Treated, Post, and the

Table 6: Falsification test and industry spillover

	Training firms	Non-training firms					
		Def. 1		Def. 2		Def. 3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Empirical strategy 1 – All firms					
Treated × Post	−3.37*	−1.53	0.87	−1.53	1.53	−2.18	3.53
	(1.79)	(1.21)	(3.58)	(1.26)	(3.73)	(2.41)	(6.46)
Treated × Post × Industry exposure			−2.40		−3.15		−6.45
			(3.95)		(4.18)		(7.92)
N	2344	8744	8744	8024	8024	2816	2816
		Empirical strategy 2 – Matched firms					
Treated × Post	−6.86***	−2.73*	0.04	−2.92**	1.18	−3.81	2.10
	(2.29)	(1.47)	(5.00)	(1.35)	(4.11)	(2.34)	(6.57)
Treated × Post × Industry exposure			−2.76		−4.14		−6.48
			(5.57)		(4.55)		(8.24)
N	1248	6112	6112	5616	5616	1824	1824

Notes: Outcome: investments per worker in €1,000. Industry exposure: Share of highly educated trainees at the detailed industry level in 1998 in %. *Def. 1:* Firms without any highly educated trainee in 1998. *Def. 2:* Firms without any highly educated trainee in 1998-2000. *Def. 3:* Firms without any trainee in 1998-2000 independent of the schooling level. Reference group: Treated × Pre. Pre: 1998–2000. Post: 2002–2004. Controlling for period and state fixed effects. Standard errors clustered at the firm level. See Appendix Figure B2.5 for the event study results.

share of highly educated trainees in an industry in 1998, while controlling for all corresponding two- and one-way interaction terms (Table 6, columns 3, 5 and 7). This analysis indeed reveals that there is no negative effect on investments among non-training firms in non-training industries. Non-training firms in training industries, in contrast, also show a non-significant drop in investments.[29]

Robustness. The negative effect on investments per worker is robust to a large range of specifications regarding data construction, the inclusion of certain states, controlling for firm fixed effects and weights. To facilitate the comparison of different results, I present difference-in-differences estimates in Figure 6, comparing the post-reform years 2002–2004 with the pre-reform years 1998–2000. I show the estimates for both the all and matched set of firms. The coefficients are consistently larger in magnitude and statistically more significantly different from zero in the matched sample.

The negative effect persists independent of the specification of the balancing requirement, i.e. when restricting to firms observed for the time period 1998 to 2004, or 1998 to 2006 instead

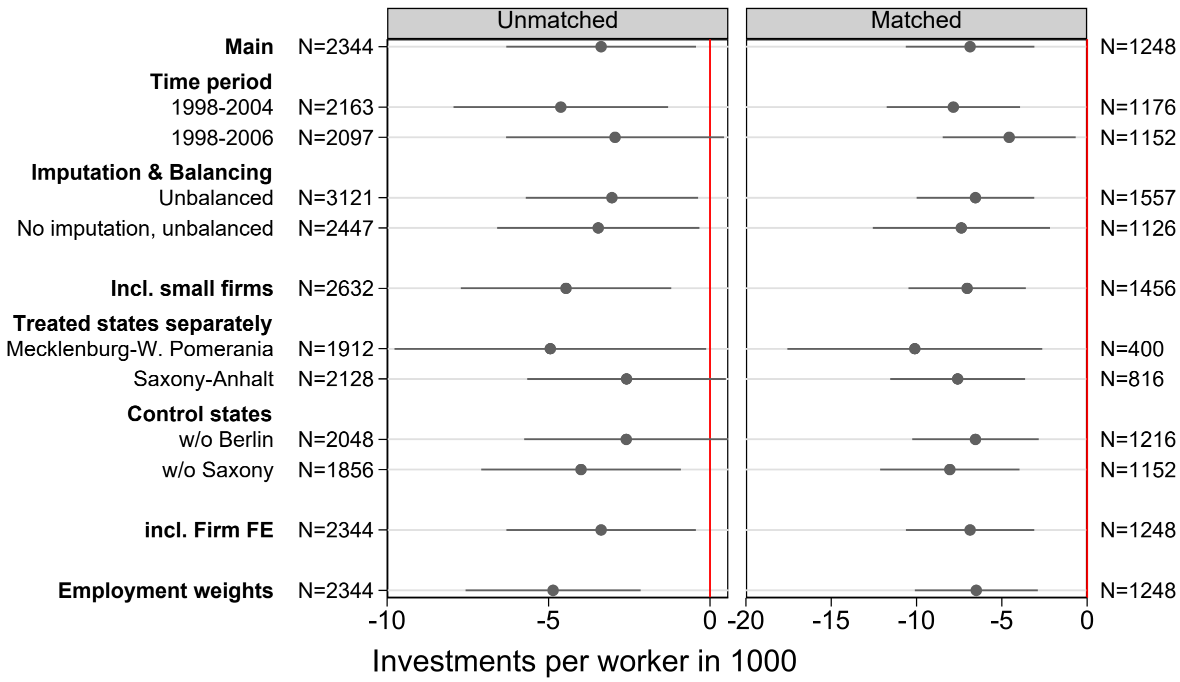
[29] Alternatively, the investment drop among non-training firms could be due to industry-state specific trends which might confound the effect of interest. If this was the case, only the triple DiD, i.e. the difference between the effect in treated training firms compared to control training firms, and the effect in treated non-training firms compared to control non-training firms, would deliver an unconfounded estimate. The triple DiD results for investments are given in Appendix Table B2.7.

of 1998 to 2005, or when fully abolishing the balancing requirement. The result is also virtually unchanged when not imputing missing values. Including firms with less than 10 employees does not alter the results significantly either.

The effect is found within both treated states separately, Mecklenburg-Western Pomerania and Saxony-Anhalt. When excluding Berlin or Saxony-Anhalt from the set of control states due to its slightly different demographic and economic trends, the result remains robust.

The coefficient is not visibly affected by the inclusion of firm fixed effects instead of state fixed effects. Weighting the observations by the firms' initial employment size in 1998 increases the negative coefficient, indicating that the impact per individual is more pronounced than the impact per firm.

Figure 6: Robustness - Investments per worker in €1,000

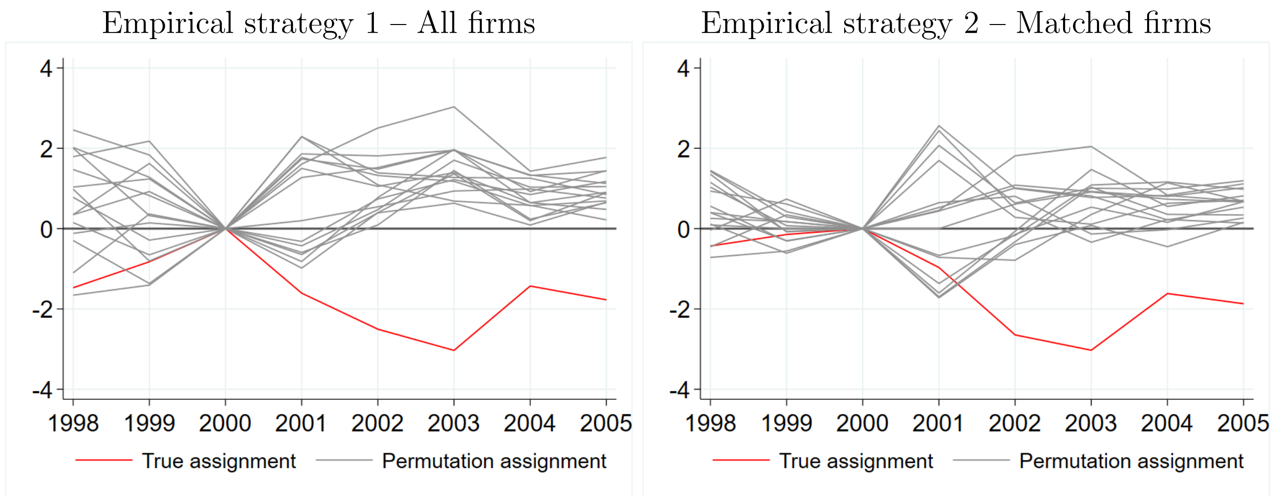


Notes: DiD estimates plus 90% confidence bands of the term Treated × Post compared to Treated × Pre plus number of observations (N). Post-reform: 2002-2004. Pre-reform: 1998-2000. Standard errors clustered at the firm level. Training firms only. *Main:* Main specification. *Time period:* requiring a balanced firm panel for 1998–2004 (1998–2006). *Unbalanced:* including firms with missing investment values. *No imputation, unbalanced:* no imputation + including firms with missing investment values. *Small firms:* Including firms with less than 10 employees. *Firm training status based on all trainee:* a training firm is a firm with at least one trainee in 1998 independent of her education. *Treated states separately:* only using one state as treated state and dropping firms from the other state. *Control states:* dropping a state from the set of control firms. *Firm FE:* including firm fixed effects instead of federal state fixed effects. *Employment weights:* Observations weighted by firms' initial employment size in 1998.

Permutation tests for inference with few clusters. Until now, I have assumed that the cluster-specific shocks are small compared to the idiosyncratic error terms at the firm level, justifying the use of standard errors clustered at the firm level. I next perform permutation tests which have been suggested a valid method for inference when the number of clusters is small (e.g. Roth et al., 2023). Figure 7 shows the t-statistics for the event study estimates based on the actual treatment assignment in red, and for all permuted treatment assignments across East German federal states in gray. The t-statistics are based on standard errors clustered at the firm level and account for sampling error of firms within states. Following the reform in 2001, the t-statistics based on the actual treatment assignment are by far more negative than any t-statistic based on a permuted treatment assignment. For periods prior to the reform, this is not the case, suggesting no differential pre-trends. Hence, the permutation test shows that it is very unlikely that cluster-level shocks only would have caused the observe investment decline. This result hold for both the sample of all firms and the sample of matched firms.

Since the number of possible permutations within East Germany is limited to 15, I repeat the permutation test across the 10 West German federal states. There was no comparable education reform in West Germany around that time. The t-statistics of the uppest and lowest 5% of the draws under permuted treatment assignment are shown in Appendix Figure B2.8. Again, the t-statistic of actual treatment assignment stands out as an outlier much smaller than the 5% and 2.5 most negative t-statistics under permuted treatment assignment.

Figure 7: Permutation test – T-statistics (Outcome: investments per worker in €1,000)



Notes: T-statistics of the event study coefficients of the actual treatment assignment (red line) and permutation assignments within East Germany (gray lines). Outcome: investments divided by total employment in 1998 in €1,000. For the corresponding figures of the regression coefficients, see Appendix Figure B2.7.

Firm-level treatment intensity – Instrumental variable regression. The average investment drop among training firms is subject to the realized distribution of trainees across training firms, and hereby subject to firms' abilities and aspirations to hire trainees despite the shortage. As a complementary analysis, I therefore instrument firms' trainee employment with a Bartik-style instrument based on firms' initial employment of highly educated trainees (i.e. exposure to the reform; share) and the reform (i.e. shift) to analyze whether training firms that suffer from larger reform-induced trainee employment reduce investments more. This analysis not only removes confounding firm selection effects; it also strengthens the argument that the investment declines are indeed caused by the negative trainee supply shock and provides an estimate of the investment decline associated with each absent highly educated trainee. I extensively discuss the identification strategy and report results in Appendix C.

The analysis reveals that more exposed firms indeed experience larger employment decreases of highly educated trainees. Likewise, firms with larger predicted employment decreases of highly educated trainees reduce investments more. In particular, each missing highly educated trainee reduces firm investments by approximately €550,000, corresponding to 9.4% of yearly average investments in training firms in 1998. This figure is lower than the one implied by the ratio between missing trainees and missing investments as identified in the event study regression above. This discrepancy might hint at spill-over effects within treated states or correlation between firm selection into trainee employment and investments: If non-investors (firms that would not have invested in absence of the supply shock) attract many trainees in face of the supply shock compared to investors (firms that would have invested in absence of the supply shock), this amplifies the average firm parameter estimated in the event study approach while not affecting the parameter identified in the IV approach.

6.2 Effect on firm technology adoption

The following section investigates whether the decrease in overall capital investments is linked to foregone technology adoption.

I next study the effect on direct indicators of firm-level technological change. Results are given in Table 7. As a first measure of firm-level technological change, I look at the technical condition of a firms' machinery (column 1). Unlike investments, technical status is a *stock* variable, expected to deteriorate as foregone investments accumulate. I therefore focus on the year 2005, when missing investments of the years 2002–2004 have accumulated. Treated training firms report an outdated technical status of their machinery compared to control training firms in 2005. The depreciation is statistically significant and meaningful in magnitude: the coefficient of -0.18 for the sample of all firms corresponds to a decrease by 5% of the average pre-reform value, and is equivalent to 18% of the firms reporting a reduction by 1 category. Since the reported technical state only changes a category in 30% of the observations from one year to the next, this corresponds to half of all firm-level technological changes. The result is comparable for the sample of matched firms. The falsification test confirms that there is no depreciation of the technical status in non-training treated firms, see Appendix Table 7.

As a second direct indicator of firm-level technological change, I study firm-level organizational change (column 2). This approach recognizes that changes in technology often accompany organizational change, such as workplace restructuring due to IT investments (Bresnahan et al., 2002). I find a substantial and statistically significant decline in organizational change among treated training firms following the reform. This decrease amounts to 0.37 (0.66 for the matched sample, respectively) reorganization measures less per firm, a drop by approximately

Table 7: Effects on firm-level technological change

	Technical status	Organizational change	Investment type (0/1)			
			Production facilities	ICT	Real estate	Transport *(Placebo)*
	(1)	(2)	(3)	(4)	(5)	(6)
			Empirical strategy 1 – All training firms			
Treated × Post	-0.18^{**}	-0.37^{**}	-0.09^{*}	-0.09^{**}	-0.08^{*}	-0.02
	(0.09)	(0.16)	(0.05)	(0.04)	(0.04)	(0.05)
	[-0.31;0.04]	[-0.70;0.19]	[-0.21;-0.02]	[-0.20;-0.11]	[-0.23;0.05]	[-0.17;0.08]
N	2341	1311	2344	2344	2344	2344
Init. outcome	3.97	1.35	0.72	0.80	0.59	0.35
			Empirical strategy 2 – Matched training firms			
Treated × Post	-0.22^{*}	-0.66^{***}	-0.09	0.00	-0.04	-0.02
	(0.13)	(0.22)	(0.07)	(0.06)	(0.07)	(0.07)
	[-0.40;0.02]	[-0.81;0.18]	[-0.14;0.00]	[-0.13;0.04]	[-0.34;0.20]	[-0.29;0.18]
N	1245	702	1248	1248	1248	1248
Init. outcome	3.98	1.41	0.71	0.79	0.58	0.33

Notes: Reference group: Treated×Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 90% confidence bands based on cluster wild t-bootstraps in square brackets. Init. outcome: Average outcome of treated firms in 1998. ICT: Information & communication technologies. For the full set of results, see Appendix Table B2.5.

one third (one half). Again, there is no comparable decline for non-training treated firms.

I next turn to the binary indicators of firm investments in certain investment categories, see columns 3–6. Unfortunately, these measures only capture the extensive margin, while I have shown above that the intensive margin is more heavily affected by the negative trainee supply shock. In the sample of all firms, treated training firms are 9 percentage points less likely to invest in production facilities and also 9% less likely to invest in ICT in the years 2002–2004 compared to untreated training firms. The investment probability remains lower in 2005. Both investment types are associated with firm-level technological change, see again Section 3. For the matched sample, the effect on investments in production facilities becomes statistically insignificant from zero, and completely fades for zero for investments in ICT. Investments in real estate, which are positively correlated to firm-level technological change, also decrease, with this effect being less pronounced. Investments in the placebo category transport, which is completely unrelated to firm-level technological change, remain unaffected. Non-training firms do not change their tendency to invest in any investment type following the reform, see Appendix Table 7.

I conclude that at least part of the investment decline is the result of reduced technology adoption. Foregone technological and organizational change may explain the reduction in internal retraining of incumbent workers established above.

Foregone investments and a slow-down in technology adoption may affect firm performance in the longer-run. However, data limitations and potentially confounding shocks pose problems when studying long-term outcomes. When looking at firm performance indicators for the years until 2005, see Appendix Table B2.6, I find no evidence of decreased sales, decreased wages, or increased firm exits. In contrast, overall firm employment is reduced, indicating that firms phase barriers to growth.

7 Economic framework and supporting evidence

7.1 Economic framework

To rationalize the complementarity between young labor market entrants and technology adoption, I next outline an economic argument. A more detailed formalization of the framework is available in Appendix D. I build up on the endogenous technological change model in Acemoglu (1998) that highlights that technology adoption will respond to price incentives set by the relative abundance of factors in the production function: If labor is scarce, increased wages incentivise the introduction labor-saving technologies and the reduction of labor-complementing technologies. I introduce capital adjustment costs to this setting. Capital adjustment costs consist of worker training in handling a new technology. As a novel key implication, technology adoption is not only endogenous to the relative abundance of factors entering the production function but also endogenous to factors entering the adjustment costs function. Since young labor market entrants have a comparative advantage in learning new skills, their absence increases capital adjustment costs and hence, hinders the adoption of technologies. In consequence, young labor market entrants are complementary to the adoption of new technologies.

Firms maximize profits by deciding whether to adopt a new, exogenously arriving and productivity-enhancing technology. New technologies might substitute or complement workers in existent tasks, while at the same time always introducing at least one new task that requires skills specific to each technology vintage. In consequence, firms incur capital adjustment costs in terms of worker training if they want to adopt the new technology. Note that this process of skill acquirement might take place within occupations but might also require changes in the firms occupational composition.

Firms can acquire skills either by retraining incumbent workers or by training young labor market entrants within a vocational training program.[30] Training costs consist of foregone production output during training and are incurred by the firms. Without training, production output of young labor market entrants is low, while incumbent workers are productive even without retraining. In consequence, firms prefer to train young labor market entrants over

[30]In principle, firms could also acquire these skills by poaching workers that have already acquired the new skills from other firms. This, however can never be a stable equilibrium. Also, it comes with other disadvantages for the firms, such as having to invest in firm-specific skills, high hiring costs, or increased risk of bad personnel decisions/less opportunities for screening.

retraining incumbent workers because their opportunity costs of training are lower and their productivity increase due to training higher.[31]

Training can only be profitable for firms if they retain workers upon training completion for a sufficient amount of time. Firms not retaining their (trained) workers will not invest in human capital of young labor market entrants and will consequently also not depend on them when it comes to the adoption of new technologies.

When young labor market entrants become temporarily unavailable, firms adopt a new technology only if the productivity gain is large enough to offset capital adjustment costs of retraining incumbent workers. If retraining incumbent workers is too costly compared to its payoff, technologies which would have been adopted if trainees were present, are not adopted. Note that this aligns with the empirical finding that internal retraining of incumbent workers does not increase in the absence of trainees.

Alternative channels. There are two alternative explanations for the complementarity between young labor market entrants and technology adoption other than their low opportunity costs and great productivity gains of learning new skills. First, according to standard human capital theory, human capital investments in young workers yield longer-term benefits in expectation (e.g. Cunha et al., 2006, or the "horizon" channel in Cavounidis & Lang, 2020). Second, young workers might generally posses more up-to-date tech skills. While both channels may play a role, they cannot fully cause the observed investment decline because they cannot explain why trainees from the previous training cohort, who are only marginally older, appear to be bad substitutes for entrants when it comes to technology adoption. The only aspect new labor market entrants are considerably different to second-year trainees is in their opportunity costs and expected payoff of acquiring new skills, as noted in Cavounidis & Lang (2020). Indeed, the Cost-Benefit Surveys of Vocational Training show that firm revenues from skilled labor activities of second-year trainees (third-year trainees) are 134% (254%) higher than of first-year trainees (Schönfeld et al., 2016, Table 18).

The reasoning above describes one potential channel leading to a reduction in investment when trainees are scarce. This mechanism requires two assumptions. First, new technologies require new skills. Second, in expectation, trainees stay at their training firm long enough to redeem firm investments in their human capital. In the next section, I provide empirical evidence in support of both assumptions, and therefore in support of this mechanism.

7.2 Empirical evidence for adjustment costs of worker training

New skills. The literature provides many examples of how new technologies require new skills, without ruling out the replacement of labor in existent tasks (e.g. Autor et al., 2003;

[31] This channel is similar to what Cavounidis & Lang (2020) call "inertia" when looking at human capital investment decisions from the worker perspective: Workers who are already specialized have higher costs of acquiring new skills.

Acemoglu & Restrepo, 2018; Deming & Noray, 2020; Autor et al., 2022). If the necessity of vintage-specific technology skills is the reason underlying firms' investment reductions, firms more exposed to skill changes should cut investments to a greater extent. Intuitively, firms with incumbent workers in occupations that have not changed recently do not rely on young labor market entrants to invest in technologies because the incumbent workers are still appropriately skilled. In contrast, firms with incumbents in occupations with recent skill changes depend on young labor market entrants to invest in new technologies because their incumbent workers do not possess the adequate skills. I measure occupational skill changes using changes in vocational training curricula from Lipowski et al. (2024). Training curricula offer an ideal measure because, first, they directly apply to the studied worker group, i.e. trainees, second, their changes are caused by technological innovation (Lipowski et al., 2024), and third, they are exogenous to individual firms since they are decided upon at the national level. I approximate firm exposure to new skills as the 1998 share of workers in one of the 18 occupations whose training curricula are updated between 1998 and 2001. There is substantial variation in firm exposure to new skills: 13% of firms are completely unexposed to skill changes; the firm at the 25th percentile of the exposure distribution employed 4% in changing occupations in 1998, and the firm at the 75th percentile 38%.

To relate the reform-induced investment drop to firm exposure to new skills, I compute the firm-level difference-in-differences for each treated firm following Schmieder et al. (2022), i.e. the difference in the investment drop 2002–2000 between a treated firm and its matched control firm:

$$\Delta\Delta \text{Inv}_j = (\text{Inv}_{j,2002} - \text{Inv}_{j,2000})_{\text{treated}} - (\text{Inv}_{j',2002} - \text{Inv}_{j',2000})_{\text{control}} \tag{2}$$

where j denotes a treated firm and j' its matched control firm. I regress this firm-level difference-in-differences on firm exposure to new skills, NewSkills.

$$\Delta\Delta \text{Inv}_j = \alpha \text{NewSkills}_j + \beta X_{jt} + u_j \tag{3}$$

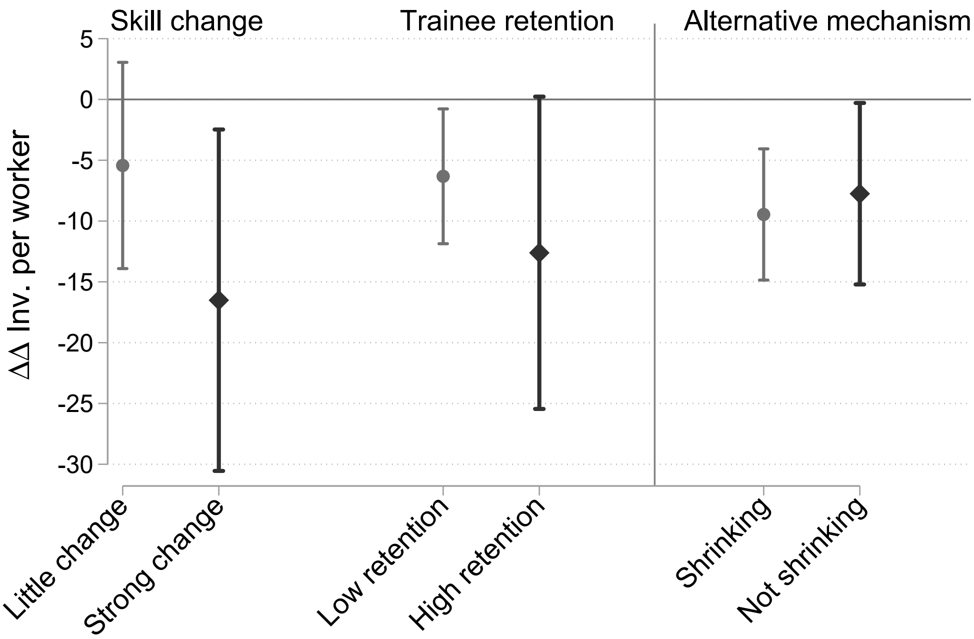
This approach is equivalent to a triple difference-in-differences specification with the triple interaction term Treated × Post × NewSkills plus all corresponding two-way and one-way interaction effects. In contrast, the specification in equation (3) is easier to interpret and allows to flexibly control for potential confounders X. In particular, I control for industry, firm size and firm exposure to the shock, i.e. number of highly educated trainees in 1998. I hence compare investment drops between two treated, equally sized firms operating in the same industry and with the same exposure to the reform, but with different exposures to new skills.[32]

The predicted investment change for firms with strong skill change (at the 90th percentile of the distribution) versus firms with little skill change (at the 10th percentile of the distribution)

[32] This implicitly assumes that the treated and and its matched control firm are similarly exposed to new skills. The balancing table (see again Appendix Table B1.2) reveals that this is approximately fulfilled.

is shown in Figure 8. In line with the idea of investment constraints of worker training, investment drops are substantially larger among firms with stronger skill changes. This finding is robust to the time period of curricula updates, see Appendix Table B3.2.[33]

Figure 8: Investment change by firm characteristics



Notes: Outcome: Change in investments per worker in €1,000 in treated training firms compared to their matched control training firms between 2002 and 2000 as defined in equation (2). Predicted drop and 90% confidence bands based on regression equation (3) using heteroscedasticity-robust standard errors. Little change and low retention correspond to the values of the 10th percentile of the corresponding distribution; strong change and high retention equal to 90th percentile. Exposure to new skills measured as the 1998 share of workers in occupations with an updated curriculum between 2000 and 2001. Trainee retention rate measured as the pre-reform share of trainees retained by the firm upon completion of the training. Controlling for industry and firm exposure to the reform (number of highly educated trainees in 1998). For robustness checks, see Appendix Tables B3.2, B3.3 and B3.4.

Trainee retention. The second assumption is that workers stay at their training firm for long enough to redeem the investments in their human capital. Indeed, the trainee retention rate in the data is high with on average approximately 40% of the trainees remaining at their training firms.However, there is variation in the retention rate across firms, see Appendix Figure B3.1. I follow Mohrenweiser & Backes-Gellner (2010) and use a firm's trainee retention rate to draw conclusions regarding the firm's training strategy.Firms with high retention rates likely see trainees as human capital investment for future production (the so-called 'investment model', see Stevens, 1994), while firms with low retention rates likely employ trainees for current production (the so-called 'production model', see Lindley, 1975). If the mechanism underlying

[33]In particular, the result is similar when looking at curricula changed directly before the missing trainee cohort and when looking at curricula changed directly following the missing trainee cohort. This shows that the investment drop is not merely a direct effect of a new curriculum, but rather the result of a general skill change in an occupation which is, among others, expressed in a curriculum change.

the reform-induced investment reduction isndeed the role of trainees as skill investment for future production, investment drops should be larger among firms with higher retention rates.

I use the same strategy as in equation (3), but with the trainee retention rate as the independent variable of interest. The retention rate is defined as the proportion of trainees staying at the firm upon training completion.[34] Figure 8, Panel B, shows the predicted investment changes for firms with high and low trainee retention rates. Consistent witht the hypothesized mechanism, firms with high retention rates reduce investments heavily in response to the reform, while treated firms with low retention rates reduce their investments much less. For the regression table including robustness checks, see Appendix Table B3.3.[35]

These two heterogeneity analyses empirically manifest that the mechanism underlying the investment drop is indeed likely to be the need for trainees to accompany the technology adoption due to their comparative advantage in skill acquisition.

Ruling out the alternative mechanism via firm growth. A shortage of young labor market entrants may also decrease firm investments because it creates an impediment to firm employment growth. If the impediment to firm employment growth is the only reason for the investment cut, only treated firms that indeed experience a net reduction in their workforce ("shrinking"), i.e. firms that do not substitute the missing trainees with other workers, should reduce investments. In contrast, treated firms that replace the missing trainees ("not shrinking") should not reduce investments. I define shrinking firms as firms with a zero or negative absolute employment growth between 2000 and 2002, and as not shrinking otherwise. Figure 8, shows that investments decline similarly strongly for shrinking and not shrinking firms. See Appendix Table B3.4 for the corresponding regression table. This finding is incompatible with the alternative mechanism via barriers to firm employment growth. It also demonstrates the unique role of young labor market entrants for firm investments: Firms may replace young labor market entrants with other workers, but these other workers are not able to fill the gap when it comes to firm technology adoption.

7.3 Relevance of capital adjustment costs

Traditional models of endogenous technological change have exclusively focused on the *incentive channel*: If wages increases, firms have an incentive to adopt labor-saving technologies, and to reduce the adoption of technologies complementing labor. In this paper, I highlight a different channel, the *constraint channel*: If employment drops, firms might not be able to adopt a new

[34] This information is based on two questions from the establishment survey on the number of trainees retained by the firm and the number of successfully completed vocational trainings. If the ratio of these two variables is not available, I construct the ratio based on the social security data. The trainee retention rate is balanced between treated training firms and control training firms, see again Appendix Table B1.2.

[35] Since the economic framework, see Appendix D, predicts a U-shaped pattern of investment decline by the retention rate, I also include the quadratic term in a further check. While the coefficients are not statistically significantly different from zero, they have the expected sign, see Appendix Table B3.3, columns 5 and 7.

technology, no matter how profitable it may be, because capital adjustment costs of worker training are prohibitively high. I now discuss the relative importance of both channels.

Let us denote $\xi^{\text{prod}} = \frac{\Delta \text{Inv}^{\text{prod}}/\text{Inv}}{\Delta w/w} \in (-\infty, +\infty)$ the elasticity of substitution between labor and technology in the production function. They are complements if $\xi^{\text{prod}} < 0$ and substitutes if $\xi^{prod} > 0$. Let us denote $\xi^{\text{adopt}} = \frac{\Delta \text{Inv}^{\text{adopt}}/\text{Inv}}{\Delta q/q} \in (0; \infty)$ the elasticity of substitution between labor and technology in the adoption process. The greater ξ^{adopt}, the more important a worker type is for the adoption of new technologies. Both channels affect investments:

$$\begin{aligned}\Delta \text{Inv}/\text{Inv} &= (\Delta \text{Inv}^{\text{prod}} + \Delta \text{Inv}^{\text{adopt}} + \Delta \text{Inv}^{\text{prod}} \Delta \text{Inv}^{\text{adopt}})/\text{Inv} \\ &= \xi^{\text{prod}} \Delta w/w + \xi^{\text{adopt}} \Delta q/q + \xi^{\text{adopt}} \xi^{\text{prod}} \Delta w/w \Delta q/q \end{aligned} \tag{4}$$

with wages w and employment q. The first addend captures the incentive channel, the second addend the constraint channel, and the third addend the interaction effect, i.e. the constraint channel on the investment change induced by the incentive channel.

A labor supply shock ΔL^S affects both employment and wages, depending on the elasticity of labor demand, the elasticity of labor supply, and labor market frictions. Equation (4) can be rewritten in the following way, with $\alpha = \frac{\Delta q/q}{\Delta L^S/L^S} \in (0; 1)$ the responsiveness of employment, and $\beta = \frac{\Delta w/w}{\Delta L^S/L^S} \in (-\infty, 0)$ the responsiveness of wages:

$$\Delta \text{Inv}/\text{Inv} = (\alpha \xi^{adopt} + \beta \xi^{prod} + \alpha \beta \xi^{\text{adopt}} \xi^{\text{prod}} \Delta L^S/L^S) \Delta L^S/L^S \tag{5}$$

Equation (5) offers a simple assessment of the effect of a labor supply shock on technology investments, depending on the type of technology (ξ^{prod}), the type of labor (ξ^{adopt}), the functioning of the labor market (α, β), and the time horizon:

- Type of technology (ξ^{prod}): If technology and labor are complements in production, i.e. $\xi^{prod} < 0$, the incentive and the constraint effect go in the same direction and investments decline when labor becomes scarce. If technology and labor are substitutes in production, i.e. $\xi^{\text{prod}} > 0$, the effect on investments is ambiguous. The more a technology acts as substitute for labor, the more likely the effect on investments is positive.

- Type of labor (ξ^{adopt}): The more important a worker type for the adoption of new technologies, i.e. the greater ξ^{adopt}, the more important the constraint channel, and the more likely investments decrease when this type of worker becomes scarce. As argued in this paper, ξ^{adopt} is larger for young labor market entrants than for incumbent workers. This parameter might be lower in other countries, given that the German vocational training system highly incentivises the transfer of modern skills.

- Functioning of the labor market (α, β): The stronger the employment response relative to the wage response, the more important the constraint channel.

- Time horizon: The longer the labor shortage persists, the more important the production channel compared to the adoption channel because firms produce in every period but adopt in the first period only.

Previous studies have identified the overall response of labor supply shocks on investments, i.e. the combination of both channels. In this paper's setting, wages do not react to the labor supply shock due to frictions. The incentive channel is therefore completely shut, i.e. $\beta = 0$, allowing to identify ξ^{adopt} for young labor market entrants: Their employment decreases by approximately 30% and investments decrease by approximately 20%, giving $\xi^{\text{adopt}} \approx 0.66$.

8 Conclusion

In this paper, I empirically demonstrate that a temporary negative supply shock of vocational trainees causally and substantially reduces firm capital investments, linked to firms' reduced adoption of technologies. The implied complementarity between young labor market entrants and firm technology adoption is explained by entrants' low opportunity costs and high expected pay-offs of skill acquisition. Firms' capital adjustment costs in terms of worker training are thus higher when young labor market entrants are scarce. Assuming that even labor-replacing technologies require some new skills, the results challenge hopes of addressing labor shortages by substituting labor with capital (e.g. Acemoglu & Restrepo, 2018), and contribute an additional dimension to macro studies predicting economic downturn in times of population aging (e.g. Jones, 2022; Kotschy & Bloom, 2023; Maestas et al., 2023). Since the time of the negative labor supply shock studied in this paper is characterized by a high unemployment rate of 18.8% (Federal Statistical Office, 2022) and an excess supply of trainees (Ministry of Education & Research, 2004), the effect may be even more pronounced in current tight labor markets.

The effects of this temporary shock likely differ from those of a long-term reduction in labor supply. The economic framework provides two key implications for the case of a long-term reduction in firm employment of young labor market entrants: First, incentives to substitute labor with capital in the production process, the channel highlighted in standard endogenous technological change models, are higher in the case of a long-term supply reduction. In the long-run, this channel may outweigh the counteracting channel of increased capital adjustment costs highlighted in this paper, potentially increasing technology adoption. Second, even if technology adoption increases in the long-run, this is associated with substantial additional (capital adjustment) costs compared to the adoption by training young labor market entrants.

The model also predicts complementarity between technology adoption and young labor market entrants in settings other than the German vocational training system. However, factors specific to the German vocational training system may cause particularly low capital adjustment cost of training vocational trainees, namely the enhanced skill transfer during vocational training due to nationally binding curricula and accompanying courses in vocational schools. Consequently, the German vocational training system serves as an effective catalyst for fostering the adoption of new technologies, as suggested by Schultheiss & Backes-Gellner (2022). At the same time, the finding that it is too expensive to retrain incumbent workers trained a couple of years ago indicates that skills acquired during vocational training are likely too specific (compare Hanushek et al., 2017).

In summary, while prior studies have shown that retraining of experienced workers is possible (Humlum et al., 2023), I provide evidence that, from the view of a profit-maximizing firm, this may not be cost-efficient. From a policy perspective, my findings therefore not only stress the importance of expanding measures to attract and mobilize young labor market entrants to foster economic growth, they also call for subsidies for retraining experienced workers.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2023). When should you adjust standard errors for clustering? The Quarterly Journal of Economics, 138(1), 1–35.

Abeliansky, A., & Prettner, K. (2017). Automation and demographic change. Available at SSRN 2959977.

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. The Quarterly Journal of Economics, 113(4), 1055–1089.

Acemoglu, D. (2002). Directed technical change. The Review of Economic Studies, 69(4), 781–809.

Acemoglu, D., & 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., & Restrepo, P. (2022). Demographics and automation. The Review of Economic Studies, 89(1), 1–44.

Adão, R., Beraja, M., & Pandalai-Nayar, N. (2020). Technological transitions with skill heterogeneity across generations (Tech. Rep.). National Bureau of Economic Research.

Ahituv, A., & Zeira, J. (2011). Technical progress and early retirement. The Economic Journal, 121(551), 171–193.

Andersson, D., Karadja, M., & Prawitz, E. (2022). Mass migration and technological change. Journal of the European Economic Association, 20(5), 1859–1896.

Angelini, D. (2023). Aging population and technology adoption (Tech. Rep.). Department of Economics, University of Konstanz.

Aubert, P., Caroli, E., & Roger, M. (2006). New technologies, organisation and age: firm-level evidence. The Economic Journal, 116(509), F73–F93.

Autor, D., Chin, C., Salomons, A. M., & Seegmiller, B. (2022). New frontiers: The origins and content of new work, 1940–2018 (Tech. Rep.). National Bureau of Economic Research.

Autor, D., & Dorn, D. (2009). This job is "getting old": measuring changes in job opportunities using occupational age structure. American Economic Review, 99(2), 45–51.

Autor, D., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. The Quarterly Journal of Economics, 118(4), 1279–1333.

Battisti, M., Dustmann, C., & Schönber, U. (2023). Technological and organizational change and the careers of workers. Journal of the European Economic Association, 21(4), 1551—1594.

Beaudry, P., Doms, M., & Lewis, E. (2010). Should the personal computer be considered a technological revolution? evidence from us metropolitan areas. Journal of Political Economy, 118(5), 988–1036.

Beicht, U., Walden, G., & Herget, H. (2004). Kosten und Nutzen der betrieblichen Berufsausbildung in Deutschland. Bielefeld: W. Bertelsmann Verlag.

Bessen, J., Goos, M., Salomons, A., & van den Berge, W. (2020). Firm-level automation: Evidence from the netherlands. In AEA papers and proceedings (Vol. 110, pp. 389–393).

Bessen, J., Goos, M., Salomons, A., & Van den Berge, W. (2023). What happens to workers at firms that automate? The Review of Economics and Statistics, 1–45.

Bibb-cost-benefit-survey 2000. gwa_1.0; reserch data center at bibb. (2009). doi: doi:10.7803/370.00.1.2.10

Bresnahan, T. F., Brynjolfsson, E., & Hitt, L. M. (2002). Information technology, workplace organization, and the demand for skilled labor: Firm-level evidence. The Quarterly Journal of Economics, 117(1), 339–376.

Büttner, B., & Thomsen, S. L. (2015). Are we spending too many years in school? Causal evidence of the impact of shortening secondary school duration. German Economic Review, 16(1), 65–86.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90(3), 414–427.

Carneiro, P., Liu, K., & Salvanes, K. G. (2022). The supply of skill and endogenous technical change: evidence from a college expansion reform. Journal of the European Economic Association.

Cavounidis, C., & Lang, K. (2020). Ben-porath meets lazear: Microfoundations for dynamic skill formation. Journal of Political Economy, 128(4), 1405–1435.

Chari, V. V., & Hopenhayn, H. (1991). Vintage human capital, growth, and the diffusion of new technology. Journal of Political Economy, 99(6), 1142–1165.

Chen, J., & Roth, J. (2023). Logs with zeros? some problems and solutions. The Quarterly Journal of Economics.

Clemens, M. A., Lewis, E. G., & Postel, H. M. (2018). Immigration restrictions as active labor market policy: Evidence from the mexican bracero exclusion. American Economic Review, 108(6), 1468–87.

Cooper, R., Haltiwanger, J., & Power, L. (1999). Machine replacement and the business cycle: lumps and bumps. American Economic Review, 89(4), 921–946.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. Handbook of the Economics of Education, 1, 697–812.

D'Acunto, F., Weber, M., & Yang, S. (2020). Manpower constraints and corporate policies. Chicago Booth Research Paper(21-03).

Danzer, A., Feuerbaum, C., & Gaessler, F. (2020). Labor supply and automation innovation. Max Planck Institute for Innovation & Competition Research Paper(20-09).

Dechezleprêtre, A., Hémous, D., Olsen, M., & Zanella, C. (2019). Automating labor: evidence from firm-level patent data. Available at SSRN 3508783.

Deming, D. J., & Noray, K. (2020). Earnings dynamics, changing job skills, and stem careers. The Quarterly Journal of Economics, 135(4), 1965–2005.

Doms, M., & Dunne, T. (1998). Capital adjustment patterns in manufacturing plants. Review of economic dynamics, 1(2), 409–429.

Dorner, M., & Görlitz, K. (2020). Training, wages and a missing school graduation cohort.

Ellguth, P., Kohaut, S., & Möller, I. (2014). The IAB Establishment Panel – Methodological essentials and data quality. Journal for Labour Market Research, 47(1–2), 27–41.

Federal Institute for Vocational Education and Training (Ed.). (2009). Datenreport zum Berufsbildungsbericht 2009. informationen und Analysen zur Entwicklung der beruflichen Bildung. (Tech. Rep.). (Table A5.4.2-1)

Federal Statistical Office, Genesis-Online. (2022a). Ausbildungsverträge: Deutschland, Jahre, Geschlecht, (November 07, 2022; data license by-2-0; own calculation/own presentation.)

Federal Statistical Office, Genesis-Online. (2022b). Bevölkerung: Bundesländer, Stichtag, Altersjahre. (November 07, 2022; data license by-2-0; own calculation/own presentation.)

Federal Statistical Office, Genesis-Online. (2022c). Studienanfänger: Deutschland, Semester, Nationalität, Geschlecht. Statistik der Studenten. (November 07, 2022; data license by-2-0; own calculation/own presentation.)

Federal Institute for Vocational Education, & Training (Eds.). (2002). Berufsbildungsbericht 2002. (Tech. Rep.). (Table 17)

Federal Institute for Vocational Education, & Training. (2022). Collectively agreed training allowances database.

Federal Ministry of Education, & Research. (2022). Tab 2.3.15. School-graduates and school-leavers, by type of school-leaving certificate, Länder and sex. (October 18, 2022; https://www.datenportal.bmbf.de/portal/en/K233.html; own calculation/own presentation.)

Federal Statistical Office, G.-O. (2022). Arbeitsmarktstatistik der Bundesagentur für Arbeit. Arbeitslose, Arbeitslosenquoten, Gemeldete Arbeitsstellen, Kurzarbeiter, Kurzarbeitende Betriebe: Deutschland/ Früheres Bundesgebiet/Neue Länder, Jahre. (November 07, 2022; data license by-2-0; own calculation/own presentation.)

Fischer, G., Janik, F., Müller, D., & Schmucker, A. (2009). The IAB Establishment Panel — Things users should know. Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften, 129(1), 133–148. doi: 10.5164/IAB.FDZD.2103.en.v1

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. Journal of Econometrics, 225(2), 254–277.

Hanushek, E. A., Schwerdt, G., Woessmann, L., & Zhang, L. (2017). General education, vocational education, and labor-market outcomes over the lifecycle. Journal of Human Resources, 52(1), 48–87.

Heine, C., Spangenberg, H., & Sommer, D. (2005). Studienberechtigte 2004: erste Schritte in Studium und Berufsausbildung; Vorauswertung der Befragung der Studienberechtigten 2004 ein halbes Jahr nach Schulabgang im Zeitvergleich. HIS.

Helbig, M., & Nikolai, R. (2015). Die Unvergleichbaren: Der Wandel der Schulsysteme in den deutschen Bundesländern seit 1949. Julius Klinkhardt.

Hobijn, B., Schoellman, T., & Vindas, A. (2019). Structural transformation by cohort (Tech. Rep.). Technical Report, Arizona State University.

Hornbeck, R., & Naidu, S. (2014). When the levee breaks: black migration and economic development in the american south. American Economic Review, 104(3), 963–90.

Humlum, A., Munch, J. R., & Jorgensen, P. (2023). Changing tracks: Human capital investment after loss of ability. University of Chicago, Becker Friedman Institute for Economics Working Paper(2023-30).

IAB establishment panel (IAB BP) — Version 9317 v1. (2019). doi: 10.5164/IAB.IABBP9317.de.en.v1

Jones, C. I. (2022). The end of economic growth? unintended consequences of a declining population. American Economic Review, 112(11), 3489–3527.

Konings, J., & Vanormelingen, S. (2015). The impact of training on productivity and wages: firm-level evidence. Review of Economics and Statistics, 97(2), 485–497.

Kotschy, R., & Bloom, D. E. (2023). Population aging and economic growth: From demographic dividend to demographic drag? (Tech. Rep.). National Bureau of Economic Research.

Kühn, S. M., Ackeren, I. v., Bellenberg, G., Reintjes, C., & Brahm, G. i. (2013). Wie viele Schuljahre bis zum Abitur? eine multiperspektivische Standortbestimmung im Kontext der aktuellen Schulzeitdebatte. Zeitschrift für Erziehungswissenschaft, 16(1), 115–136.

Landesamt, H. S. (Ed.). (2023). Erwerbstätigenrechnung. Erwerbstätige in den Ländern der Bundesrepublik Deutschland 1991 bis 2022. Berechnungsstand: Mai 2023 (Tech. Rep.).

Le Barbanchon, T., Ronchi, M., & Sauvagnat, J. (2023). Hiring frictions and firms' growth. Available at SSRN 4105264.

Lerche, A. (2019). Investment tax credits and the response of firms.

Lewis, E. (2011). Immigration, skill mix, and capital skill complementarity. The Quarterly Journal of Economics, 126(2), 1029–1069.

Lightcast. (2021). The demographic drought: The approaching labor shortage.

Lindley, R. M. (1975). The demand for apprentice recruits by the engineering industry, 1951-71. Scottish Journal of Political Economy, 22(1), 1–24.

Linked-employer-employee-data from the IAB (LIAB): LIAB-cross-sectional model 2 (LIAB QM2) 1993–2019, version 1. (2021). doi: doi:10.5164/IAB.LIABQM29319.de.en.v1

Lipowski, C., Salomons, A., & Zierahn-Weilage, U. (2024). Keeping up with the computers – how vocational training responds to new technology. (Unpublished Manuscript)

Liu, Y., & Mao, J. (2019). How do tax incentives affect investment and productivity? firm-level evidence from china. American Economic Journal: Economic Policy, 11(3), 261–291.

MacDonald, G., & Weisbach, M. S. (2004). The economics of has-beens. Journal of Political Economy, 112(S1), S289–S310.

Maestas, N., Mullen, K. J., & Powell, D. (2023). The effect of population aging on economic growth, the labor force, and productivity. American Economic Journal: Macroeconomics, 15(2), 306–332.

Marcus, J., & Zambre, V. (2019). The effect of increasing education efficiency on university enrollment evidence from administrative data and an unusual schooling reform in germany. Journal of Human Resources, 54(2), 468–502.

Ministry of Education, F., & Research (Eds.). (2004). Berufsbildungsbericht 2003. (Tech. Rep.).

Mohrenweiser, J., & Backes-Gellner, U. (2010). Apprenticeship training: for investment or substitution? International Journal of Manpower.

Morin, L.-P. (2015). Cohort size and youth earnings: evidence from a quasi-experiment. Labour Economics, 32, 99–111.

Muehlemann, S., Dietrich, H., Pfann, G., & Pfeifer, H. (2022). Supply shocks in the market for apprenticeship training. Economics of Education Review, 86, 102197.

Neuber-Pohl, C., Pregaldini, D., Backes-Gellner, U., Dummert, S., & Pfeifer, H. (2023). How negative labor supply shocks affect training in firms: Lessons from opening the swiss-german border. Swiss Leading House "Economics of Education" Working Paper, 203.

OECD. (2023). Retaining talent at all ages. Retrieved from https://www.oecd-ilibrary.org/content/publication/00dbdd06-en doi: https://doi.org/https://doi.org/10.1787/00dbdd06-en

Roth, J. (2022). Pretest with caution: Event-study estimates after testing for parallel trends. American Economic Review: Insights, 4(3), 305–322.

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. Journal of Econometrics.

Ruf, K., Schmidtlein, L., Seth, S., Stüber, H., & Umkehrer, M. (2021). Linked-employer-employee-data from the IAB (LIAB): LIAB-cross-sectional model 2 (LIAB QM2) 1993–2019. FDZ-Datenreport, 03/2021 (en). (Tech. Rep.). doi: 10.5164/IAB.FDZD.2103.en.v1

San, S. (2023). Labor supply and directed technical change: Evidence from the termination of the bracero program in 1964. American Economic Journal: Applied Economics, 15(1), 136–63.

Sauvagnat, J., & Schivardi, F. (2023, 02). Are executives in short supply? Evidence from death events. The Review of Economic Studies, rdad027. Retrieved from https://doi.org/10.1093/restud/rdad027 doi: 10.1093/restud/rdad027

Schmieder, J. F., von Wachter, T. M., & Heining, J. (2022). The costs of job displacement over the business cycle and its sources: evidence from Germany (Tech. Rep.). National Bureau of Economic Research.

Schönfeld, G., Jansen, A., Wenzelmann, F., & Pfeifer, H. (2016). Kosten und Nutzen der dualen Ausbildung aus Sicht der Betriebe. Ergebnisse der fünften BIBB-Kosten-Nutzen-Erhebung (Tech. Rep.).

Schultheiss, T., & Backes-Gellner, U. (2022). Does updating education curricula accelerate technology adoption in the workplace? evidence from dual vocational education and training curricula in switzerland. The Journal of Technology Transfer, 1–45.

Socio-Economic Panel. (SOEP, 2019). Socio-Economic Panel (SOEP), Version 34, Data for years 1984-2017 (SOEP-Core v34). (DOI: 10.5684/soep.v34)

Stevens, M. (1994). An investment model for the supply of training by employers. The Economic Journal, 104(424), 556–570.

Wolter, S. C., & Ryan, P. (2011). Apprenticeship. In Handbook of the Economics of Education (Vol. 3, pp. 521–576). Elsevier.

Zeira, J. (1998). Workers, machines, and economic growth. The Quarterly Journal of Economics, 113(4), 1091–1117.

Zwick, E., & Mahon, J. (2017). Tax policy and heterogeneous investment behavior. American Economic Review, 107(1), 217–48.

A Data

A.1 Data imputation

The data imputation procedure consists of two steps. Table A1.1 shows the number of observations and firms after each imputation step. Variables from the administrative dataset also need imputation since they are not filled whenever the firm has no valid interview. In a first step, I linearly interpolate missing values in up to two consecutive years if the firm has valid entries before and afterwards. I interpolate starting variables, such as total investments and total employment, and compute variables building on them based on their interpolated values, such as investments per worker, or the indicator for large investments. For binary variables, e.g. investment type, I impute a zero if the linear interpolation is a non-integer number.

In a second step, I constantly extrapolate values at the start (1998, 1999) and at the end of the observation window (2004, 2005) for firms known to have existed in these years based on information from the social security records.

Table A1.2 shows how imputation and balancing affects observations and firms. Imputed observations are not significantly different from non-imputed observations, except with respect to total investments (column 2). Imputed investments tend to be smaller, likely because imputing investment spikes (see below) is barely feasible. The imputation procedure successfully recovers small firms with smaller investments which otherwise would have been lost due to the balancing requirement, enhancing the representativeness of the sample (column 4). In general, balanced firms are larger and have more investments, even after imputation (column 5). I therefore compute robustness checks which confirm the results in the non-imputed and/or unbalanced dataset.

Table A1.1: Imputation steps

	Initial dataset	After imputation				
		Interpolation		Extrapolation		Combined
		All firms				
Number of observations with non-missing values for...						
... # highly educated trainees	10,344	10,444	+1.0%	11,088	+6.2%	+7.2%
... investments	9,896	10,101	+2.1%	11,088	+9.8%	+12.0%
Number of balanced firms	670	757	+13.0%	1,386	+83.1%	+106.9%
		Training firms				
Number of observations with non-missing values for...						
... highly educated trainees	2,227	2,250	+1.0%	2,344	+4.2%	+5.3%
... investments	2,140	2,182	+2.0%	2,344	+7.4%	+9.5%
Number of balanced firms	168	193	+14.9%	293	+51.8%	+74.4%

Notes: Numbers refer to the (restricted and balanced) sample ultimately used in the subsequent analyses. For years without a valid interview, information from the administrative employment data is also missing and has to be imputed.

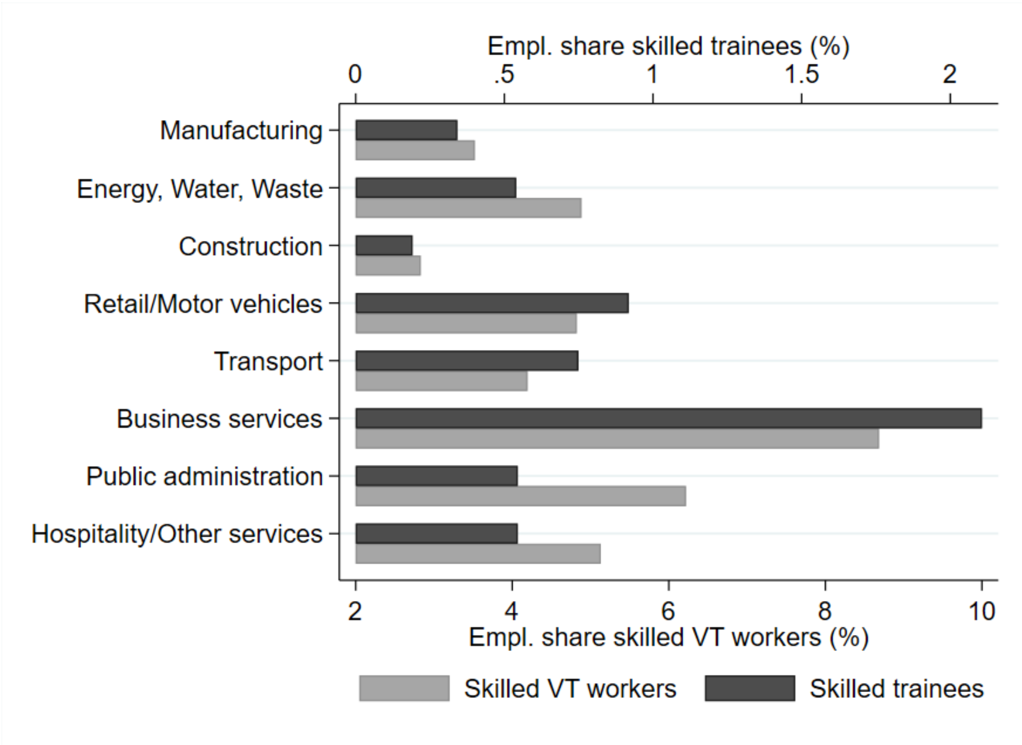
Table A1.2: Descriptives – Imputation and balancing

	Unbalanced		Balanced		Δ Balanced
	Non-imputed	Non-imp. - Imputed	Non-imputed	Non-imp. - Imputed	Unbalanced - Balanced
	(1)	(2)	(3)	(4)	(5)
# workers	136.7	3.4	167.6	19.6***	−14.7***
% highly educ. trainees	0.69	0.01	0.66	0.05	0.07***
Inv. per worker	12.8	1.04***	15.1	2.3***	−0.99***
			Industry		
Agriculture	0.04	0.00	0.07	0.02***	−0.01***
Manufacturing	0.34	−0.01**	0.24	−0.07***	0.03***
Energy, water, waste	0.03	0.00	0.03	0.00	0.00
Construction	0.12	0.00	0.09	0.00	0.02***
Retail/motor vehicles	0.09	0.00	0.09	0.00	0.00
Transport	0.03	0.00	0.05	0.01***	0.00
Business services	0.14	0.00	0.14	0.01**	0.01***
Public administration	0.15	0.00	0.21	0.03***	−0.04***
Other services	0.07	0.00	0.07	0.00	−0.01**

Notes: Unbalanced: All firms. Balanced: Only firm with non-missing investments for 1998–2005. Δ Balanced: Difference between the average in the imputed unbalanced dataset and the average in the imputed balanced dataset. Significance stars for the two-sided t-test of the difference. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

A.2 Descriptives and summary statistics

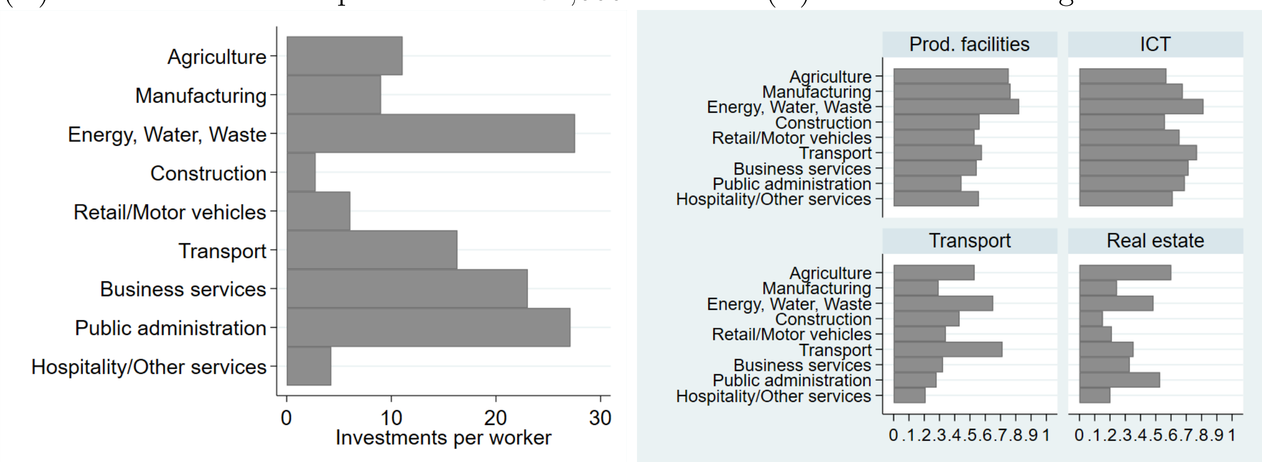
Figure A2.1: Highly educated trainees by industry



Notes: Share of highly educated trainees (highly educated VT workers) among total firm employment. Observations at the firm-year level. VT=Vocational training.

Figure A2.2: Investments by industry

(A) - Mean investments per worker in €1,000 (B) - Investment categories



Notes: Firm-year level observations. *Panel B:* Share of observations with investments in the specified investment category.

Table A2.1: Investment and technology indicators in the establishment panel

Variable	Survey Question	Manipulation	Frequency
Inv. per worker	What was the approximate sum of all investments in t?	Divided by number of workers in 1998 from the administrative records and excluding outliers	Yearly
Expansion inv.s per worker	What share of total investments made was attributed to the expansion of your establishment?	Multiplied by investments per worker	Yearly
Inv. in ICT	Did your establishment invest in one or more of the following areas in the last business year of t? EDP, information and communication technology?		Yearly
Inv. in production facilities	Did your establishment invest in one or more of the following areas in the last business year of t? Production facilities, plant and equipment, furniture and fixture?		Yearly
Inv. in transport	Did your establishment invest in one or more of the following areas in the last business year of t? Means of transport, transportation systems?		Yearly
Inv. in real estate	Did your establishment invest in one or more of the following areas in the last business year of t? Real estate and buildings?		Yearly
Technology status of machinery	How do you assess the overall technical state of the plant and machinery, furnitures and fixtures of this establishment compared to other establishments in the same industry? "1" - state-of-the-art equipment. "5" - completely out-of-date.	Inverted order	Yearly except for 2004
Organizational Change	Has one or more of the following organisational changes been carried out within your establishment/office in the last two years? (1) Restructuring of departments or areas of activities, (2) Downward shifting of responsibilities and decisions, (3) Introduction of team work/ working groups with their own responsibilities, (4) Introduction of units/departments carrying out their own cost and result calculations	Sum of the four	1998, 2000, 2001, 2004, 2007, 2010, 2012, 2014, 2015, 2017

Table A2.2: Sum of total investments

Share of obs. w/o investments	19.3%
5^{th} percentile	€10,000
25^{th} percentile	€61,224
50^{th} percentile	€331,633
75^{th} percentile	€2,200,000
95^{th} percentile	€15,077,000
Mean	€2,679,418

Notes: For the corresponding table for total investments, see Figure 2, Panel A.

B Additional results

B.1 Additional results – Identification strategy

Figure B1.1: Employment of workers with completed vocational training/university studies

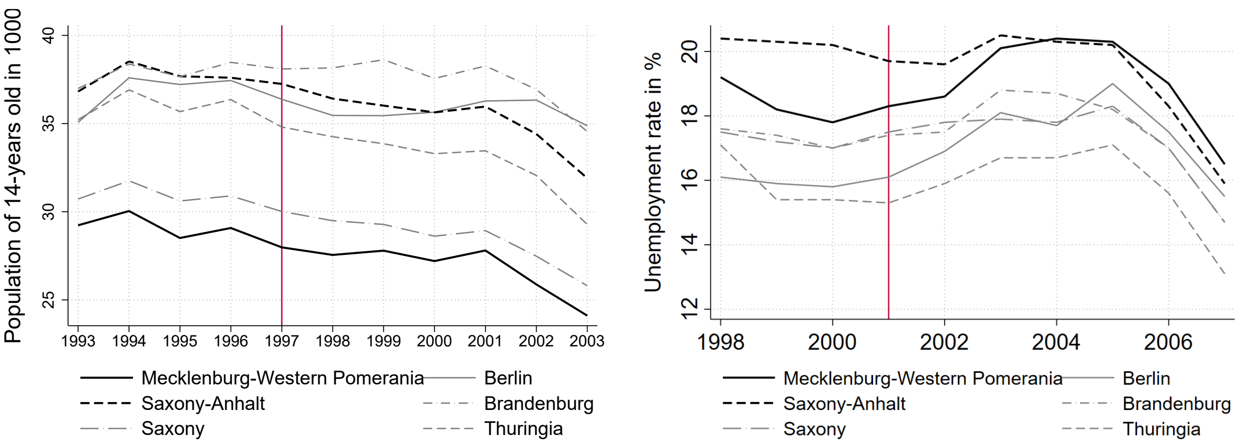
(A) Highly educated VT employment per firm (B) TE employment per firm

Notes: LIAB, own calculations. VT – Vocational training, TE – tertiary education. Red vertical solid line: Last year before the reform-induced trainee supply shock. Red vertical dashed line in Panel A: Last year before the reform-induced shock of workers with completed vocational training arrives (conditional on starting training in 2001 and taking three years). Including trainees of the dual system only. Red vertical dashed line in Panel B: Last year before the reform-induced supply shock of tertiary educated workers arrives (conditional on starting university in 2001 and taking five years).

Figure B1.2: Demographic and economic trends across federal states

(A) Number of 14-years-old by state

(B) Unemployment rate by state



Notes: Panel A: Source: Federal Statistical Office, Genesis-Online (2022b). The number for Saxony is divided by two for better visibility. *Panel B:* Source: Federal Statistical Office, Genesis-Online (2022b).

Table B1.1: Imbalance before and after matching – Targeted variables

	Training firms			Non-training firms		
	Mean Treated	Δ Mean Unmatched	Δ Mean Matched	Mean Treated	Δ Mean Unmatched	Δ Mean Matched
			Industry			
Agriculture	0.04	0.03	0	0.06	0.00	0
		(1.49)	0		(0.15)	0
Manufacturing	0.21	-0.11*	0	0.26	-0.10***	0
		(-1.91)	0		(-3.36)	0
Energy, Water, Waste	0.05	0.01	0	0.03	0.01	0
		(0.25)	0		(0.71)	0
Construction	0.05	-0.01	0	0.10	0.01	0
		(-0.41)	0		(0.70)	0
Retail/Motor vehicles	0.06	-0.03	0	0.11	0.03*	0
		(-0.84)	0		(1.69)	0
Transport	0.02	-0.01	0	0.04	0.01	0
		(-0.39)	0		(0.58)	0
Business services	0.21	0.03	0	0.12	0.01	0
		(0.49)	0		(0.47)	0
Public administration	0.31	0.08	0	0.18	0.03	0
		(1.31)	0		(1.16)	0
Hospitality/Other services	0.05	0.01	0	0.08	-0.00	0
		(0.42)	0		(-0.18)	0
			Mahalanobis matching variables			
% highly educated trainees 2000	2.32	0.31	0.50	0.12	-0.02	-0.01
		(0.73)	(0.96)		(-0.40)	(-0.13)
% highly educated trainees 1999	2.51	0.21	0.41	0.04	-0.01	-0.00
		(0.47)	(0.75)		(-0.38)	(-0.01)
% highly educated trainees 1998	2.69	0.05	0.29	0.00	0.00	0.00
		(0.10)	(0.51)			
Investment per worker 2000	18.28	0.44	3.04	13.52	-0.47	2.24
		(0.12)	(0.77)		(-0.24)	(1.04)
Investment per worker 1999	17.76	-1.36	2.67	14.34	-0.85	2.19
		(-0.39)	(0.74)		(-0.44)	(1.03)
Investment per worker 1998	17.43	-3.13	1.84	14.87	-0.36	2.05
		(-0.83)	(0.46)		(-0.18)	(0.91)
Pre avg. log(employment)	5.16	0.00	-0.29*	4.12	0.08	0.02
		(0.01)	(-1.69)		(1.15)	(0.31)
N		293	156		1093	764

Notes: Δ Mean: Mean Treated - Mean Control; N: Number of firms. T-statistic of the two-sided t-test of the difference in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B1.2: Imbalance before and after matching – Untargeted variables

	Training firms			Non-training firms		
	Mean Treated	Δ Mean Unmatched	Δ Mean Matched	Mean Treated	Δ Mean Unmatched	Δ Mean Matched
Pre avg. prob to invest	0.92	-0.01	0.00	0.83	-0.04**	-0.02
		(-0.19)	(0.11)		(-2.21)	(-0.92)
Pre avg. prob for large investments	0.38	0.06	0.03	0.22	-0.01	0.03
		(1.10)	(0.48)		(-0.36)	(1.27)
Pre avg. inv in prod facilities	0.73	-0.06	-0.04	0.60	-0.05**	-0.03
		(-1.24)	(-0.70)		(-2.12)	(-1.08)
Pre avg. inv in ICT	0.85	0.02	-0.00	0.65	-0.04	-0.03
		(0.59)	(-0.10)		(-1.47)	(-1.09)
Pre avg. inv in real estate	0.57	0.06	-0.06	0.33	-0.01	0.03
		(0.97)	(-0.93)		(-0.36)	(1.22)
Pre avg. inv in transport	0.36	-0.07	-0.16**	0.37	0.01	0.05*
		(-1.33)	(-2.43)		(0.48)	(1.79)
Pre avg. org Change	1.20	0.13	0.13	0.68	-0.05	-0.02
		(0.92)	(0.77)		(-1.00)	(-0.41)
Pre avg. tech status	3.98	0.05	0.00	3.78	-0.09**	-0.04
		(0.59)	(0.04)		(-2.10)	(-0.83)
Pre avg. trainee retention rate	0.60	0.04	0.02	0.50	-0.05**	-0.05*
		(1.09)	(0.43)		(-2.43)	(-1.88)
Pre avg. rate of skill change	29.68	5.18	2.72	24.20	-1.14	1.48
		(1.22)	(0.52)		(-0.62)	(0.73)
N		293	156		1093	764

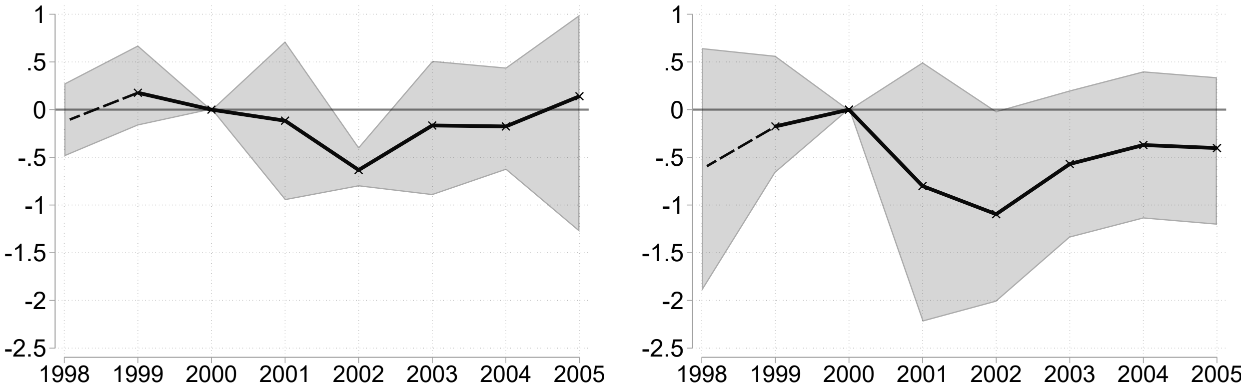
Notes: Δ Mean: Mean Treated - Mean Control; N: Number of firms. T-statistic of the two-sided t-test of the difference in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B.2 Additional results – Estimation results

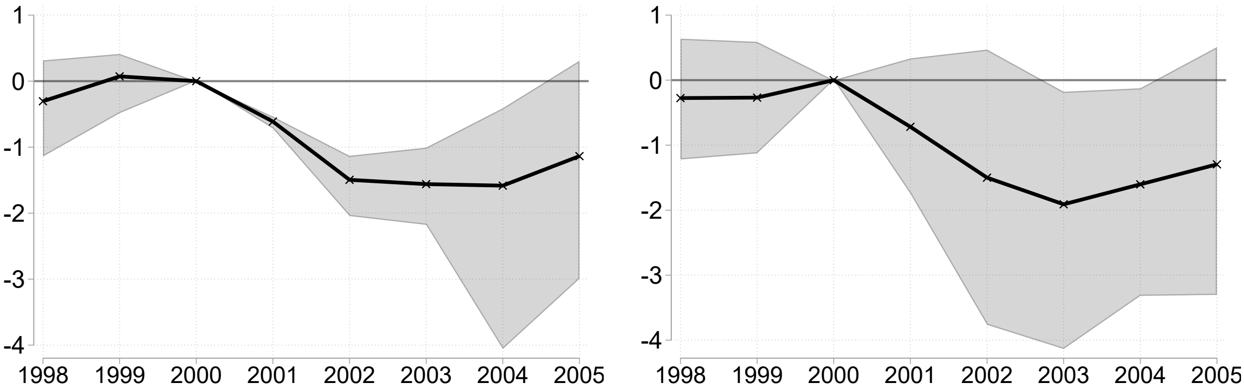
Figure B2.1: Bite of the shock – Cluster wild t-bootstrap confidence intervals

(A) New highly educated trainee hires

Empirical strategy 1 – All firms | Empirical strategy 2 – Matched firms

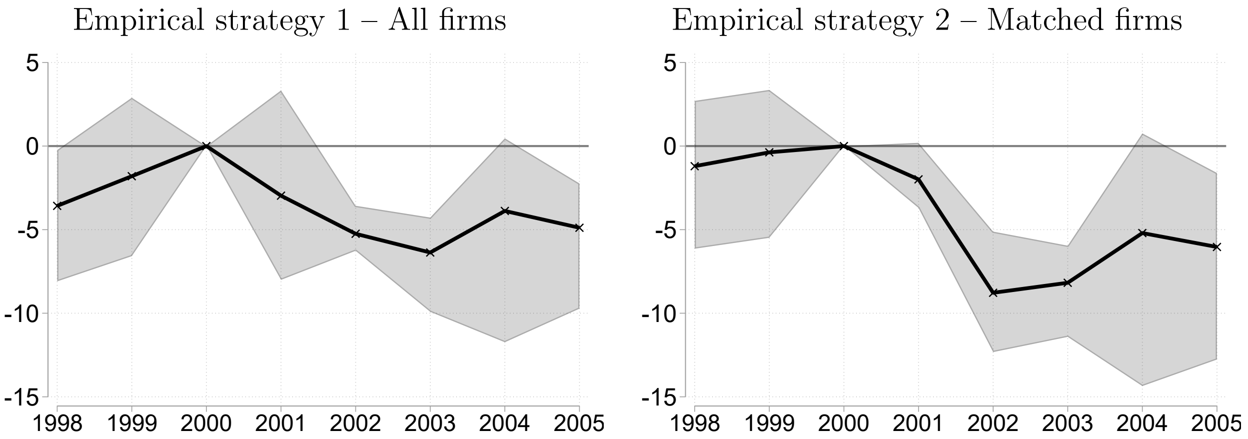


(B) Number of highly educated trainees



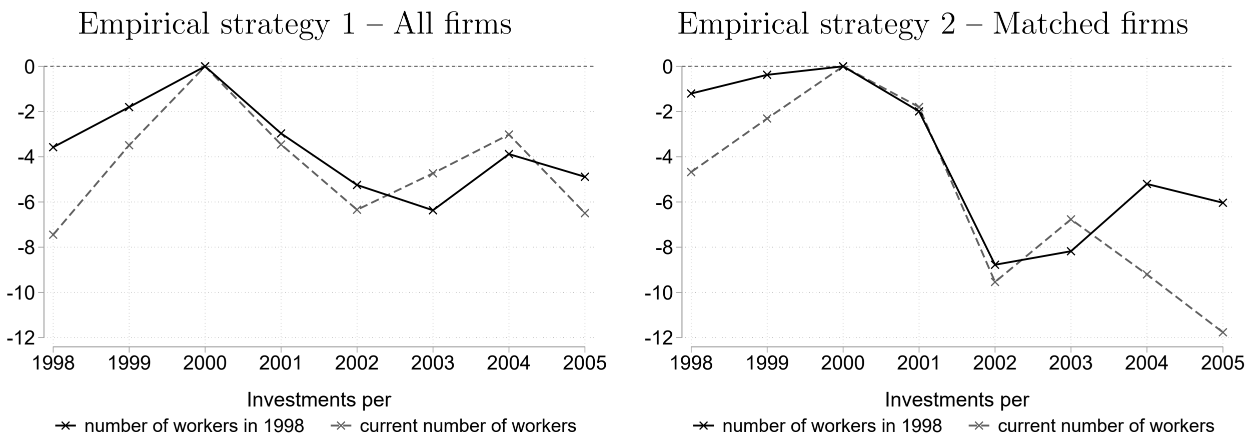
Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. 90% Confidence intervals based on cluster wild t-bootstraps following Cameron et al. (2008). Hirings in 1998 should be taken with caution.

Figure B2.2: Investments per worker in €1,000 - Wild cluster t-bootstrap confidence intervals



Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Outcome: investments divided by total employment in 1998 in €1,000. 90% Confidence intervals based on cluster wild t-bootstraps following Cameron et al. (2008).

Figure B2.3: Investments per worker in €1,000



Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Standard errors clustered at the firm level.

Table B2.1: DiD Results - Wage and substitution effects (Full table)

	Wage effects		Substitution effects			
	Log wages highly educ. trainees		# low-educ. trainee hires	# highly educ. commuting trainee hires	Internal retraining	Log VT employment
	(1)	(2)	(3)	(4)	(5)	(6)
			Empirical strategy 1 – All training firms			
Treated × Roll-out	−0.02	−0.05*	0.16	0.05	−0.01	−0.05*
	(0.04)	(0.03)	(0.71)	(0.06)	(0.05)	(0.03)
Treated × Post	−0.03	−0.04	−0.42	0.03	−0.09*	−0.13**
	(0.04)	(0.03)	(0.84)	(0.05)	(0.05)	(0.05)
Treated × Phase-out	−0.08*	−0.07*	0.64	−0.04	−0.14**	−0.16
	(0.04)	(0.04)	(1.05)	(0.10)	(0.06)	(0.10)
N	1758	1758	2295	2018	2227	2344
Firm FE		X				
Init. outcome	3.00	3.00	6.33	0.04	0.42	4.87
			Empirical strategy 2 – Matched training firms			
Treated × Roll-out	0.00	−0.03	0.15	0.04	−0.05	−0.04
	(0.04)	(0.04)	(1.03)	(0.09)	(0.06)	(0.04)
Treated × Post	0.01	−0.01	−0.04	0.04	−0.07	−0.07
	(0.05)	(0.04)	(1.11)	(0.06)	(0.07)	(0.07)
Treated × Phase-out	−0.05	−0.04	−0.04	0.08	−0.12	−0.03
	(0.05)	(0.04)	(1.51)	(0.11)	(0.11)	(0.14)
N	908	908	1233	1082	1190	1248
Firm FE		X				
Init. outcome	3.00	3.00	6.00	0.03	0.43	4.82

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. Column 4: a commuter is defined as a person living and working in two distinct federal states. This variable is available from 1999 onward. Column 5: Internal retraining is the sum of retraining incidences at the firm-year level. VT: completed vocational training.

Table B2.2: DiD Results - Substitution effects continued

	Trainee retention rate	# VT separations	# VT hires	# low-educ. VT hires	# highly educ. VT hires
	(1)	(2)	(3)	(4)	(5)
			Empirical strategy 1 – All training firms		
Treated × Post	-0.09^{**}	4.93	3.28	3.35	-0.08
	(0.04)	(3.88)	(2.87)	(2.74)	(0.28)
	[-0.03;0.00]	[-11.33;25.35]	[-6.39;17.47]	[-6.42;15.93]	[-0.08;1.65]
N	2260	2281	2295	2295	2295
Init. outcome	0.39	23.51	14.04	12.40	1.64
			Empirical strategy 2 – Matched training firms		
Treated × Post	-0.10^{**}	5.45	4.43	4.44	-0.01
	(0.04)	(7.43)	(3.54)	(3.36)	(0.34)
	[-0.06;0.09]	[-3.45;25.93]	[3.86;12.35]	[2.17;3.31]	[-0.17;1.40]
N	1215	1224	1233	1233	1233
Init. outcome	0.39	21.26	14.05	12.38	1.68

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. Column 1: The trainee retention rate is equal to the share of trainees (independent of school education) which are offered a working contract after training graduation. VT: completed vocational training. For the main results on wage and substitution effects, see Table 4.

Table B2.3: DiD Results – Investments (Full table)

	Investments per worker		Intensive vs. extensive margin			
	per init. # of workers	per current # of workers	Any inv. (0/1)	Log(Inv.)	Large inv. (1/0)	Combined
	(1)	(2)	(3)	(4)	(5)	(6)
	Empirical strategy 1 – All training firms					
Treated × Roll-out	−1.17	0.19	0.01	−0.29	0.00	−0.20
	(1.58)	(2.42)	(0.04)	(0.21)	(0.05)	(0.26)
Treated × Post	−3.37*	−1.05	−0.05	−0.33**	−0.07	−0.65***
	(1.79)	(2.39)	(0.03)	(0.16)	(0.04)	(0.24)
Treated × Phase-out	−3.09	−2.85	−0.07	−0.12	−0.02	−0.61*
	(2.60)	(3.54)	(0.05)	(0.21)	(0.07)	(0.35)
% of init. outcome	-19%	-6%	-6%	-4%	-13%	-10%
N	2344	2344	2344	2069	2069	2344
Init. outcome	17.43	17.43	0.89	7.45	0.54	6.62
	Empirical strategy 2 – Matched training firms					
Treated × Roll-out	−1.47	0.53	0.01	−0.41*	−0.03	−0.29
	(1.90)	(2.71)	(0.05)	(0.24)	(0.07)	(0.33)
Treated × Post	−6.86***	−6.17*	−0.03	−0.50**	−0.12**	−0.55
	(2.29)	(3.46)	(0.05)	(0.21)	(0.05)	(0.37)
Treated × Phase-out	−5.51*	−9.44	−0.02	−0.40	−0.10	−0.37
	(2.82)	(5.98)	(0.07)	(0.29)	(0.08)	(0.52)
% of init. outcome	-39%	-35%	-3%	-7%	-22%	-8%
N	2344	2344	2344	2069	2069	2344
Init. outcome	17.68	17.68	0.88	7.41	0.55	6.55

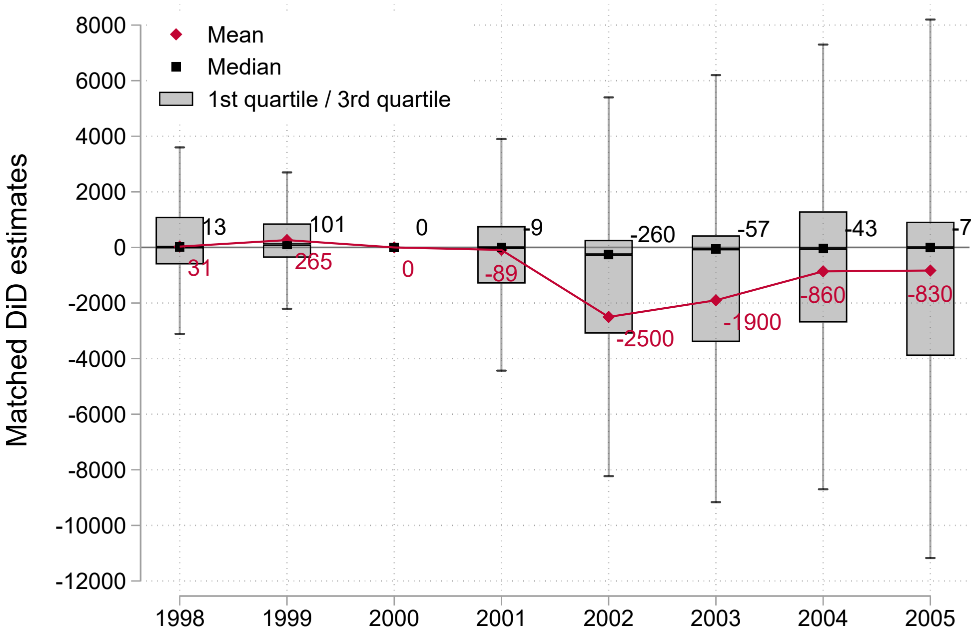
Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. For the corresponding event study figure for investments per worker, see Figure 4.

Table B2.4: DiD Results – Investment effects continued

	Investments per worker			Large investments		
	Overall	Business serv. + Public admin.	Manufacturing	Uppest tercile	Uppest decile	Uppest industry-specific tercile
	(1)	(2)	(3)	(4)	(5)	(6)
	Empirical strategy 1 – All training firms					
Treated × Post	-3.37^{*}	-1.86	-3.62^{**}	-0.07	-0.02	-0.07
	(1.79)	(3.45)	(1.39)	(0.04)	(0.02)	(0.04)
	[-7.18;-0.13]	[-11.20;4.05]	[-0.82;3.21]	[-0.16;0.08]	[-0.08;-0.08]	[-0.21;0.06]
N	2344	1040	808	2069	2069	2069
Init. outcome	17.43	24.22	6.49	0.54	0.08	0.35
	Empirical strategy 2 – Matched training firms					
Treated × Post	-6.86^{***}	-10.52^{**}	-3.00	-0.12^{**}	-0.07^{**}	-0.12^{**}
	(2.29)	(4.04)	(1.88)	(0.05)	(0.03)	(0.05)
	[-10.70;-1.72]	[-20.55;-2.86]	[-1.20;7.71]	[-0.21;-0.01]	[-0.09;0.07]	[-0.21;0.06]
N	1248	672	336	1068	1068	1068
Init. outcome	17.68	24.22	6.49	0.55	0.09	0.35

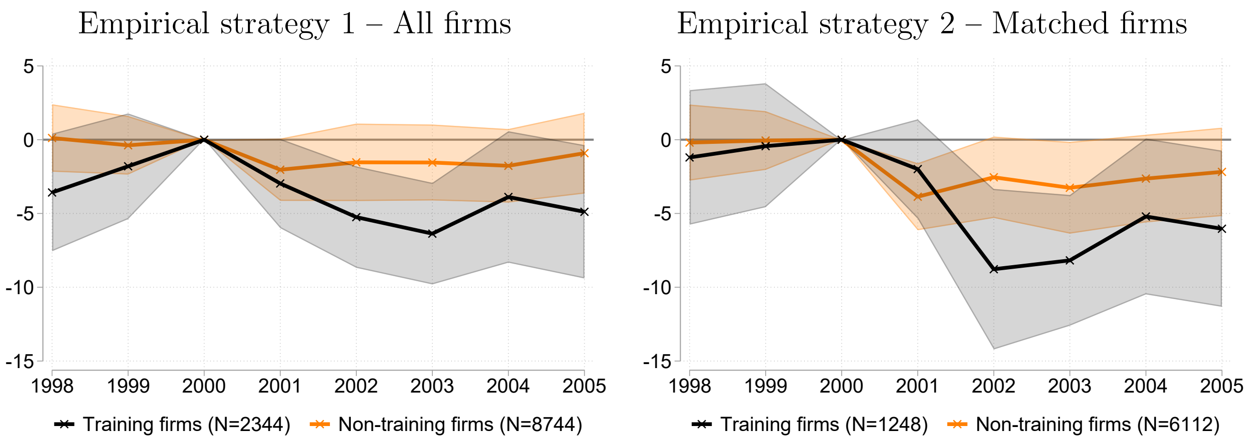
Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. Column 4: a commuter is defined as a person living and working in two distinct federal states. This variable is available from 1999 onward only. Column 5: Internal retraining is the sum of individual retraining incidences at the firm-year level. VT: completed vocational training.

Figure B2.4: Distribution of matched DiD estimates – Total investments



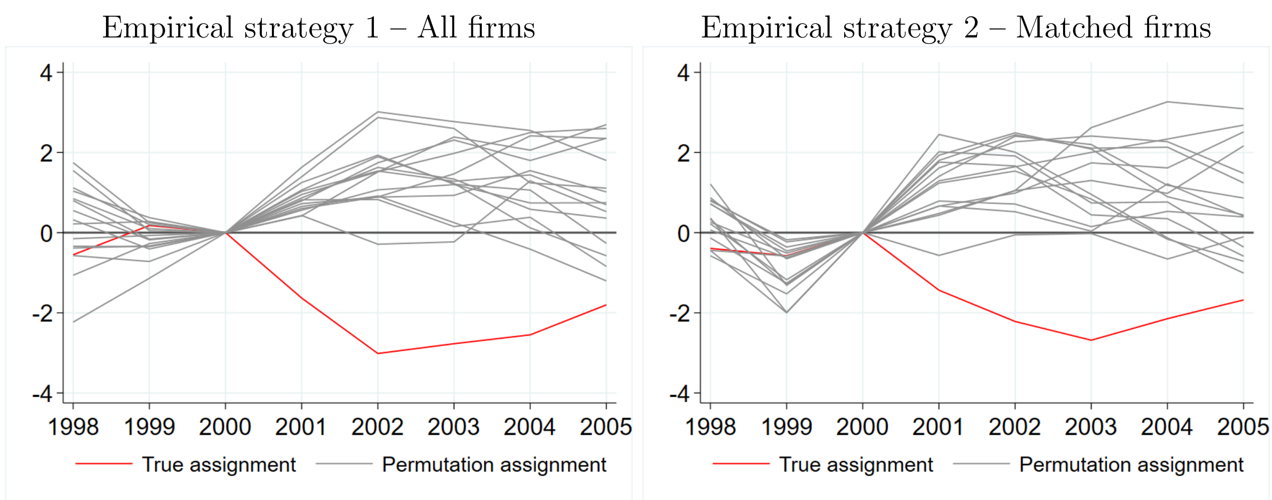
Notes: Distribution of the matched firm-level difference-in-differences in investments per worker. *Red:* Average. *Black:* Median. *Box:* 25th and 75th percentile. *Adjacent values:* 25th percentile-1.5*the interquartile range (75th percentile +1.5*the interquartile range). For the corresponding difference-in-differences estimate see Table 5, column 4.

Figure B2.5: Investments per worker in €1,000 in non-training firms



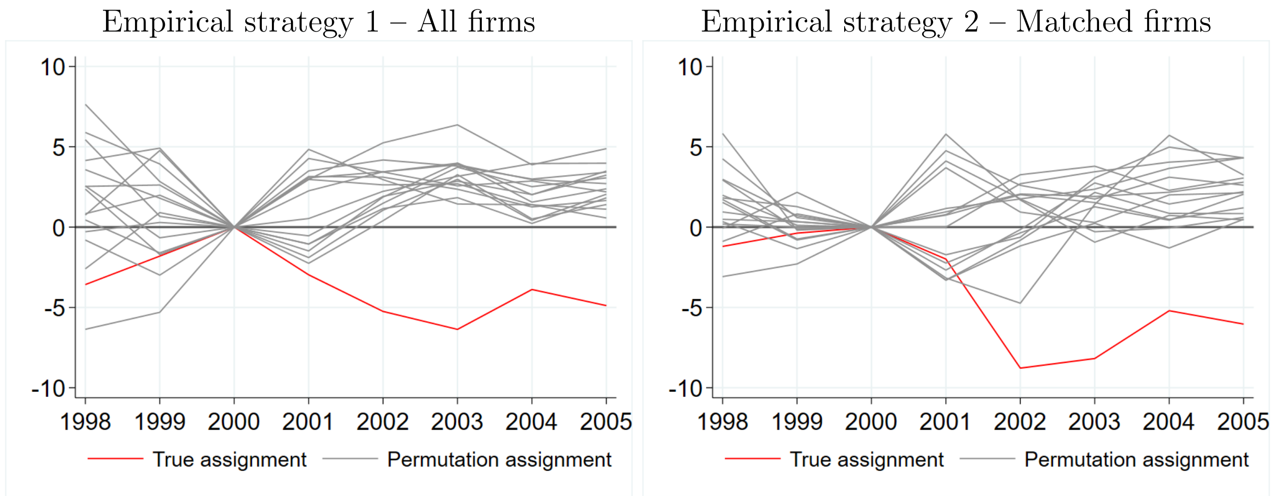
Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Standard errors clustered at the firm level. Outcome: investments divided by total employment in 1998 in €1,000. Training firms: Firms with at least one highly educated trainee in 1998. Non-training firms: Firms with no highly educated trainee in 1998.. For the corresponding difference-in-differences estimate see Table 5.

Figure B2.6: Permutation test – T-statistics (Outcome: Highly educated trainee employment)



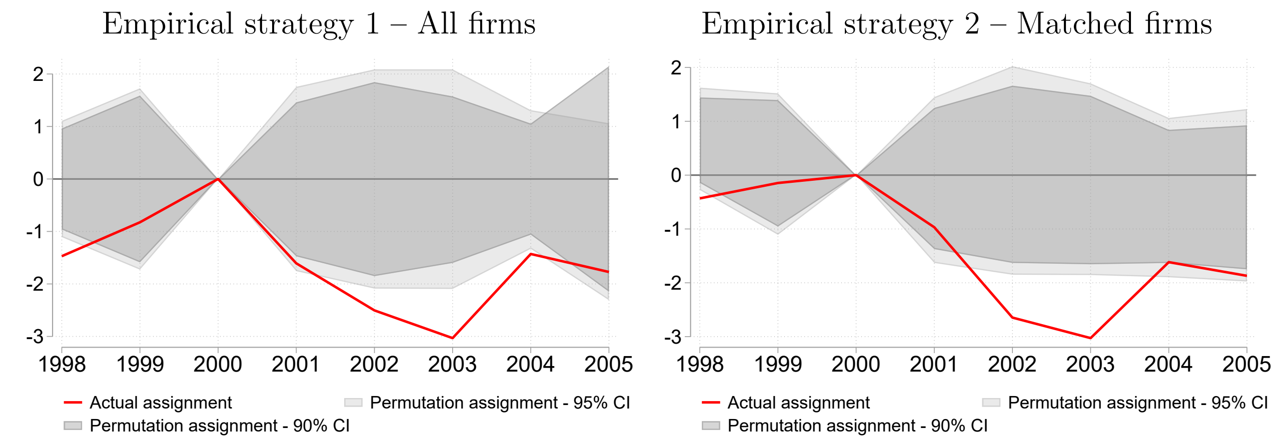
Notes: T-statistics of the event study coefficients of the actual treatment assignment (red line) and all potential permutation assignments within East Germany (gray lines). Outcome: Number of highly eductaed trainees per firm. For the corresponding figures of the t-statistics, see Figure 7.

Figure B2.7: Permutation test – Estimation coefficients (Outcome: investments per worker in €1,000)



Notes: T-statistics of the event study coefficients of the actual treatment assignment (red line) and all potential permutation assignments within East Germany (gray lines). Outcome: investments divided by total employment in 1998 in €1,000. For the corresponding figures of the t-statistics, see Figure 7.

Figure B2.8: Permutation test West Germany – T-statistics (Outcome: investments per worker in €1,000)



Notes: T-statistics of the event study coefficients of the actual treatment assignment (red line) and permutation assignments within West Germany (gray lines). Outcome: investments divided by total employment in 1998 in €1,000.

Table B2.5: DiD Results - Effects on firm-level technological change (Full table)

	Technical status	Organizational change	Investment type (0/1) Production facilities	ICT	Real estate	Transport *(Placebo)*
	(1)	(2)	(3)	(4)	(5)	(6)
			Empirical strategy 1 – All training firms			
Treated × Roll-out	−0.03	−0.10	−0.02	−0.06	−0.09*	−0.04
	(0.07)	(0.14)	(0.06)	(0.05)	(0.05)	(0.06)
Treated × Post	−0.18**	−0.37**	−0.09*	−0.09**	−0.08*	−0.02
	(0.09)	(0.16)	(0.05)	(0.04)	(0.04)	(0.05)
Treated × Phase-out		−0.27	−0.11*	−0.14**	−0.05	−0.07
		(0.17)	(0.06)	(0.06)	(0.05)	(0.06)
	[-0.31;0.04]	[-0.70;0.19]	[-0.21;-0.02]	[-0.20;-0.11]	[-0.23;0.05]	[-0.17;0.08]
N	2341	1311	2344	2344	2344	2344
Init. outcome	3.97	1.35	0.72	0.80	0.59	0.35
			Empirical strategy 2 – Matched training firms			
Treated × Roll-out	−0.03	−0.09	−0.05	−0.05	−0.09	0.00
	(0.10)	(0.22)	(0.07)	(0.06)	(0.06)	(0.07)
Treated × Post	−0.22*	−0.66***	−0.09	0.00	−0.04	−0.02
	(0.13)	(0.22)	(0.07)	(0.06)	(0.07)	(0.07)
Treated × Phase-out		−0.20	−0.06	−0.03	−0.01	−0.05
		(0.24)	(0.09)	(0.07)	(0.08)	(0.10)
	[-0.40;0.02]	[-0.81;0.18]	[-0.14;0.00]	[-0.13;0.04]	[-0.34;0.20]	[-0.29;0.18]
N	1245	702	1248	1248	1248	1248
Init. outcome	3.98	1.41	0.71	0.79	0.58	0.33

Notes: Full set of results for Table 7. Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. ICT: Information & communication technologies.

Table B2.6: DiD Results - Other outcomes

	Sales per worker (1)	Log employment (2)	Log wages (3)
	Empirical strategy 1 – All training firms		
Treated × Post	-2.79	-0.11^{**}	0.00
	(26.06)	(0.05)	(0.01)
	[-90.31;77.69]	[-0.21;0.06]	[-0.03;-0.03]
N	1260	2344	2344
Init. outcome	234.83	5.21	4.17
	Empirical strategy 2 – Matched training firms		
Treated × Post	-2.79	-0.11^{**}	0.00
	(26.06)	(0.05)	(0.01)
	[-122.30;118.61]	[-0.17;0.17]	[-0.03;0.05]
N	558	1248	1248
Init. outcome	245.37	5.18	4.17

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998.

Table B2.7: Triple DiD Results - Investments

	Investments per worker		Intensive vs. extensive margin			
	per init. # of workers	per current # of workers	Any inv. (0/1)	Log(Inv.)	Large inv. (1/0)	Combined
	(1)	(2)	(3)	(4)	(5)	(6)
	Empirical strategy 1 – All firms					
Treated × Post × Training	-1.84	-0.08	0.04	-0.27	-0.04	-0.54^{**}
	(2.15)	(2.69)	(0.04)	(0.18)	(0.05)	(0.27)
	[-5.65;0.98]	[-4.83;-0.83]	[-0.01;0.18]	[-0.43;0.29]	[-0.19;0.11]	
% of init. outcome	-11%	-0%	+ 35%	-4%	-8%	-8%
N	11088	11088	11088	8950	8950	11088
Init. outcome	17.43	17.43	0.11	7.45	0.54	6.62
	Empirical strategy 2 – Matched firms					
Treated × Post × Training	-4.13	-3.03	0.00	-0.27	-0.05	-0.25
	(2.71)	(3.77)	(0.06)	(0.23)	(0.06)	(0.40)
	[-7.90;0.55]	[-16.80;1.88]	[-0.09;0.13]	[-0.68;0.17]	[-0.18;0.08]	[-1.06;0.42]
% of init. outcome	-23%	-17%	-3%	-4%	-10%	-4%
N	11088	11088	11088	8950	8950	11088
Init. outcome	17.68	17.68	0.12	7.41	0.55	6.55

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998.

B.3 Additional results – Mechanism

Table B3.1: Trainee retention rate and firm investments

	Investments per worker				Technical status of machinery			
	Retention – Def. 1		Retention – Def. 2		Retention – Def. 1		Retention – Def. 2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Trainee retention rate	4.02***	3.90***	3.32**	3.73***	0.08**	0.09**	0.06*	0.06*
	(1.38)	(1.33)	(1.47)	(1.39)	(0.03)	0.03)	(0.03)	(0.03)
Observations	2903	2903	2580	2580	2903	2903	2580	2580
Mean outcome	15.55	15.55	15.55	15.55	3.86	3.86	3.86	3.86
Controls		✓		✓		✓		✓

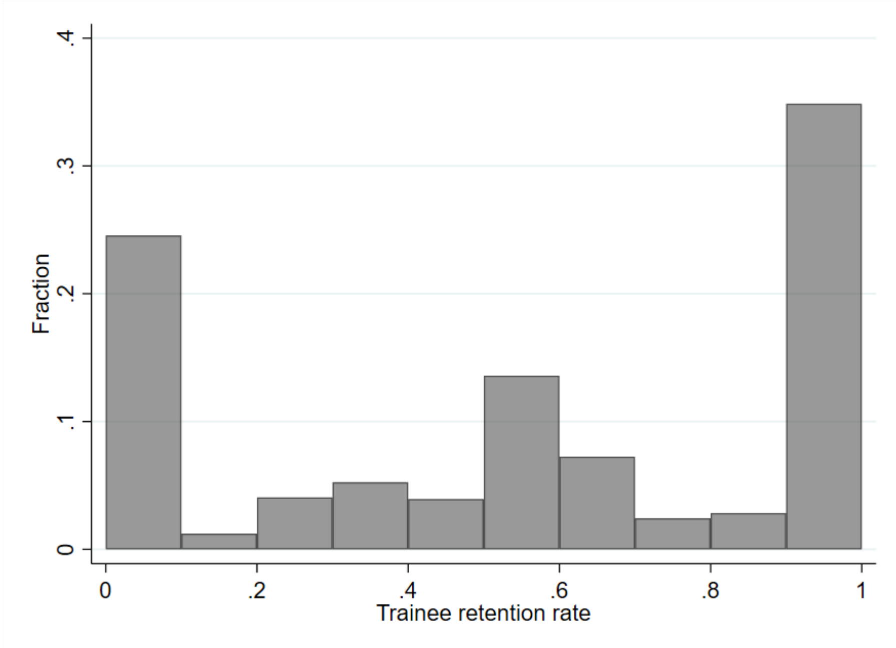
Notes: Outcome: Column 1–4: investments per worker in €1,000. Columns 5–8: Technical status of machinery on a scale from 1 "completely-out-of-date" to 5 "state-of-the-art". To avoid confusion by the education reform, including firms in control states only. Controls include year fixed effects, industry, firm size categories and federal states. Firm-year observations with at least one trainee. Definition 1: Based on the survey question "How many of the newly qualified apprentices are being offered a permanent position?". If missing, filled with the share of retained trainees from the administrative data. Definition 2: Share of retained trainees after training graduation based on the administrative data. Trainee retention rate is lagged by three years because investment decisions in their human capital are made when training starts. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B3.2: Heterogeneity by arrival rate of new skills

	(1)	(2)	(3)	(4)	(5)	(6)
New skills	0.11	-0.09	-0.09	-0.08	-0.10	-0.13
	(0.08)	(0.15)	(0.15)	(0.15)	(0.13)	(0.13)
Industry		✓	✓	✓	✓	✓
Exposure			✓	✓	✓	✓
Firm size				✓	✓	✓
Definition of new skills based on						
... Years	1998-2001	1998-2001	1998-2001	1998-2001	1996-1999	2001-2003
Matched DiD Observations	78	78	78	78	78	78
Underlying total observations	312	312	312	312	312	312

Notes: Outcome: Change in investments per worker in treated training firms compared to their matched control training firms between 2002 and 2000, see equation (2). New skills: Exposure to new skills measured as the 1998 share of all workers (trainees) in occupations with an updated curriculum between 2000 and 2001 (1998–2001). In 2000/2001, 18 occupations got updated, between 1998 and 2001, 33 occupations got updated. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For the corresponding main figure, see Figure 8.

Figure B3.1: Distribution of the trainee retention rate



Notes: Firm-year level observations. Histogram of the trainee retention rate, based on the question in the establishment panel and supplemented with information from the administrative data. A trainee is counted as retained (= 1), if she was observed as worker with vocational training at the same firm the year following her apprenticeship, and 0 otherwise.

Table B3.3: Heterogeneity by trainee retention rate

	Retention rate - Def. 1					Retention rate - Def. 2	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Trainee retention rate	-1.66	-6.78	-6.81	-6.29	-20.03	-20.86	-18.15
	(9.67)	(10.68)	(10.69)	(10.11)	(22.75)	(11.27)	(24.18)
Trainee retention rate^2					15.18		16.86
					(30.14)		(31.28)
Industry		✓	✓	✓	✓	✓	✓
Exposure			✓	✓	✓	✓	✓
Firm size				✓	✓	✓	✓
Matched DiD Observations	78	78	78	78	78	77	77
Underlying total observations	312	312	312	312	312	308	308

Notes: Outcome: Change in investments per worker in treated training firms compared to their matched control training firms between 2002 and 2000, see equation (2). Trainee retention rate measured as the pre-reform share of trainees retained by the firm upon completion of the training. Definition 1: Based on the survey question "How many of the newly qualified apprentices are being offered a permanent position?". If missing, filled with the share of retained trainees from the administrative data. Definition 2: Share of retained trainees after training graduation based on the administrative data. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For the corresponding main figure, see Figure 8.

Table B3.4: Heterogeneity by firm employment growth

	(1)	(2)	(3)	(4)
Not shrinking	3.10	1.80	1.71	1.40
	(5.42)	(5.65)	(5.83)	(6.09)
Industry		✓	✓	✓
Exposure			✓	✓
Firm size				✓
Matched DiD Observations	78	78	78	78
Underlying total observations	312	312	312	312

Notes: Outcome: Change in investments per worker in treated training firms compared to their matched control training firms between 2002 and 2000, see equation (2). Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For the corresponding main figure, see Figure 8.

C Instrumental variable regression

I examine the treatment effect along the intensive margin of the negative trainee supply shock using a complementary identification strategy.

I first run an OLS regression of firm investments on firm employment of highly educated trainees controlling for firm and year fixed effects. Firm fixed effects absorb potentially confounding factors such as firm size and industry. The result is shown in Table C1, column 1. One additional highly educated trainee is associated with €78,600 of investments, though this coefficient is not statistically different from zero.

The OLS approach has two main drawbacks: First, it generally mixes trainee supply and trainee demand. Second, it is based on the realized distribution of trainees across firms in face of the supply shock, which depends on the firms' capabilities and aspirations to employ trainees when they are scarce which are potentially related to their investor status. To disentangle trainee supply from demand, and to abstract from endogenous factors affecting the distribution of trainees among firms, I next estimate the following two-stage-least-squares (2SLS) model:

$$\mathrm{Inv}_{jbt} = N^{\mathrm{Trainee}}_{jbt} + \psi_t + \pi_j + \epsilon_{jt} \tag{C1}$$

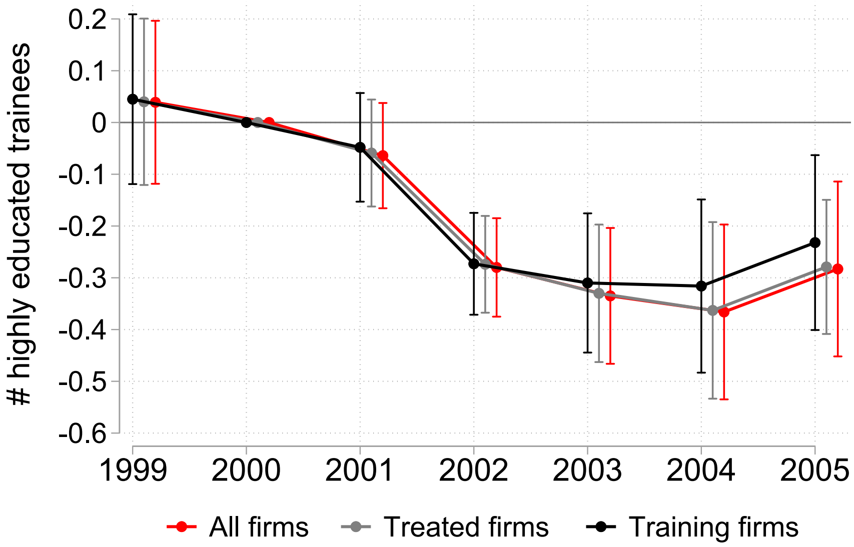
$$N^{\mathrm{Trainee}}_{jbt} = \sum_{t=1999, t\neq 2000}^{2005} \gamma_t (N^{\mathrm{Trainee}}_{j,1998} \times \mathrm{Treated}_{b(j)} \times \mathrm{Year}_t) + \psi_t + \pi_j + \epsilon_{jbt} \tag{C2}$$

with equation (C2) capturing the first-stage equation and equation (C1) the second stage. Inv denotes investments, N^{Trainee} employment of highly educated trainees, j firm, b federal states, and t calendar year. Treated is a binary variable with Treated $= 1$ if the firm is located in a state undergoing the education reform and zero otherwise. ψ_t captures year fixed effects, and ϕ_j firm fixed effects.

The first stage predicts firm-level supply of highly educated trainees based on firms' initial employment levels of highly educated trainees, $N^{\mathrm{Trainee}}_{1998}$, comparable to the *shares* in a shift-share IV, and the binary incidence of a reform-induced trainee supply shock at the state level, Treated $\times$ Year, comparable to the *shifts*. The exogeneity of the instruments stems from the random assignment of the trainee supply shock, i.e. the education reform, to firms and years. Since employment of highly educated trainees in 1998 is expected to directly impact investments of the same year, I run the regression for the years from 1999 onward.

Figure C1, Panel A, shows the coefficients of interest, γ_t, of the first stage (red line), namely the effect of one highly educated trainee more in 1998 in a treated state. Supporting the relevance of the instrument, employment of highly educated trainees drops significantly more for firms with high initial trainee employment. In particular, each additional trainee prior to the reform is associated with 0.3–0.4 trainees less in 2002–2004. This is almost perfectly in line with the event study estimates in Section 5. With an F-statistic of 22.24, the instruments are relevant.

Figure C1: IV results – First stage



Notes: Outcome: Number of highly educated trainees. Coefficients plus 90% confidence intervals of the term $N^{\text{Trainee}}_{j,1998} \times \text{Treated}_{b(j)} \times \text{Year}_t$ in equation (C2). Standard errors clustered at the firm level.

The result of the second stage is displayed in Table C1, column 2. Trainee supply and investments are significantly positively associated. In particular, one trainee less reduces investment by €572,000, corresponding to 9.4% of yearly average investments of training firms in 1998, or 20.7% of yearly average investments of all firms. This IV estimate is substantially larger than the OLS estimate. This suggests that the distribution of trainees across treated training firms is indeed correlated with firms' investor status. In particular, firms that would have invested in absence of the shock employ less trainees during the shortage compared to firms that would not have invested in absence of the shock.

The instrument is equal to zero for both firms in control states and firms in treated states with no exposure to the shock. This is equivalent to assuming that trainee employment in control states and in unexposed firms evolved in the same way. To relax this assumption, I rerun the regression among treated firms only (gray line in Figure C1 and columns 3 and 4 in Table C1) and among training firms only (black line; columns 5 and 6). Convincingly, both the results from the first stages and the results from the second stages look very similar.

In summary, this complementary identification strategy confirms the negative impact of reduced trainee supply on firm investments. It demonstrates that firms more affected by the negative trainee supply shock reduce investments more: each highly educated trainee supplied less corresponds to approximately €550,000 investments less. This figure is lower than the one implied by the ratio between missing trainees and missing investments as identified in the event study regression above. This discrepancy might hint at spill-over effects within treated states or correlation between firm selection into trainee employment and investments: If non-investors (firms that would not have invested in absence of the supply shock) attract many

Table C1: IV results – Second stage

	All firms		Treated firms		Training firms	
	OLS	2SLS	OLS	2SLS	OLS	2SLS
	(1)	(2)	(3)	(4)	(5)	(6)
N^{Trainee}	78.6	572.7***	228.0*	551.9**	87.1	600.9**
	(51.2)	(217.0)	(121.6)	(218.2)	(55.4)	(236.1)
N	9702	9702	3241	3241	2051	2051
p–value KP		0.026		0.020		0.062
F–Stat		22.20		23.30		14.50

Notes: Outcome: Total investments in €1,000. 1999-2005. N^{Trainee}: number of highly educated trainees. F-Stat gives the robust Kleibergen-Paap Wald rk F statistic. P-value KP gives the p-value of the Kleibergen-Paap test, producing valid inference for the second stage coefficient even when instruments are weak. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

trainees in face of the supply shock compared to investors (firms that would have invested in absence of the supply shock), this amplifies the average firm parameter estimated in the event study approach while not affecting the parameter identified in the IV approach.

D Economic framework

In this Appendix, I provide a formal exposition of the economic reasoning why labor market entrants act as complements to the adoption of new technologies. In essence, I introduce capital adjustment costs to a simplified version of the endogenous technological change model in Acemoglu (1998). Capital adjustment costs consist of worker training in handling a new technology and vary by worker groups. As a novel key implication, this set-up makes technology adoption endogenous to the relative abundance of factors entering the adjustment cost function.

Baseline setting. Suppose that firms operate and employees work in overlapping generations for two periods $T = 2$. This assumption is relaxed to an infinite time horizon below. In each period t, each firm j produces one final good Y using labor L and production technologies τ with fixed marginal productivities A_τ under the following production function:

$$Y_{jt} = \sum_{\tau=0}^{\mathcal{T}} y_{j\tau} = \sum_{\tau=0}^{\mathcal{T}} A_\tau L_{jt\tau} \tag{D1}$$

For simplicity, assume that the intermediate outputs or tasks produced with the different technologies, $y_{j\tau}$, are perfect substitutes.[36] As in Acemoglu (1998), technologies require skills, i.e. only workers trained for a specific technology, L_τ, can handle this technology. This production function zooms in on the labor reinstatement channel of new technologies, i.e. the aspect that new technologies create new tasks performed by humans, and abstracts from the automation channel, i.e. the aspect that new technologies automate tasks previously performed by humans. Including the automation aspect introduces a counteracting channel which I discuss below. Except for workers' ability to handle technologies, workers are homogeneous. The price for the final product is fixed to one for simplicity.

At the beginning of each period, a unit-sized cohort of homogeneous, untrained workers, L_0, with a baseline productivity A_0 enters the labor market, and a new technology τ becomes exogenously available. Compared to the previous technology $\tau-1$, the new technology increases worker productivity by $\Delta A_\tau = A_\tau - A_{\tau-1}$. ΔA_τ follows a Poisson distribution with a rate of 1, $\Delta A_\tau \sim Pois(1)$. Hence, technological progress is always positive, but rarely large.

Firms decide whether to adopt the new technology at the start of the period in order to maximize profits. In order to adopt the new technology, firms (re-)train a fraction Ψ_{τ_0} of workers of each initial productivity type. Training uniformly takes one period across technologies and workers.[37] Since workers within a cohort are homogeneous, firms always either retrain all or no worker of one entry cohort, $\Psi_{\tau_0} = \{0; 1\}$, such that worker cohorts and worker productivity

[36] This assumption can be relaxed and the production function rewritten as a CES production function. The advantage of assuming perfect substitutability between intermediate outputs is that the additive separability of the intermediate outputs allows to target *changes* in firms profits instead of total firm profits below in the firms maximization problem below.

[37] Allowing for shorter training periods of entrants compared to incumbents due to more up-to-date technical skills would present an additional factor why training entrants is more profitable than training incumbents.

types coincide. Wages w_τ are in proportion to, but below worker productivity due to firms' monopsony power, $w_\tau = \theta A_\tau$ with $\theta \in (0,1)$. Benefits from technology-induced productivity increases are hence not completely passed on to workers.[38] For simplicity, assume that workers do not switch firms.

As a novel and decisive feature of this framework, costs of technology adoption consist of capital adjustment costs C equal to costs of worker training born by the firm. For simplicity, I assume that there are no other capital costs. Training costs are equal to the sum of foregone outputs of all workers undergoing training in period $t = 1$:

$$C_{j\tau} = \sum_{\tau_0=0}^{\mathcal{T}-1} A_{\tau_0} \Psi_{\tau_0} L_{jt\tau_0} \tag{D2}$$

Firm maximization problem. Given the irreversibility of (human) capital investments and the additive separability of intermediate outputs in the production function, firms maximize profits by maximizing additional profits from adopting the new technology by deciding whether or not to (re-)train each initial worker type. Additional profits are equal to the net surplus in output minus wages in period $t = 2$ minus capital adjustment costs:

$$\max_{\sum_{\tau_0=0}^{\mathcal{T}-1} \Psi_{\tau_0}} \Delta Y_{j\tau} - \Delta W_{j\tau} - C_{j\tau} \tag{D3}$$

The net surplus in output minus wages is equal to the sum of productivity increases minus wage increases across all initial worker types L_{τ_0} trained in the new technology:

$$\Delta Y_{j\tau} - \Delta W_{j\tau} = (1-\theta) \sum_{\tau_0=0}^{\mathcal{T}-1} \Psi_{\tau_0} L_{jt\tau_0} (A_\tau - A_{\tau_0}) \tag{D4}$$

The profitability of (re-)training hence decreases in workers' initial productivities: The net output surplus of training is decreasing in workers' initial productivity (equation (D4)), while training costs are increasing in workers' initial productivity (equation (D2)). Combining equations (D1)–(D4), it follows that firms train a worker type L_{τ_0} as long as additional profits exceed additional costs, i.e. as long as the following condition between the productivity of the new technology, A_τ, and initial productivity, A_{τ_0}, holds:

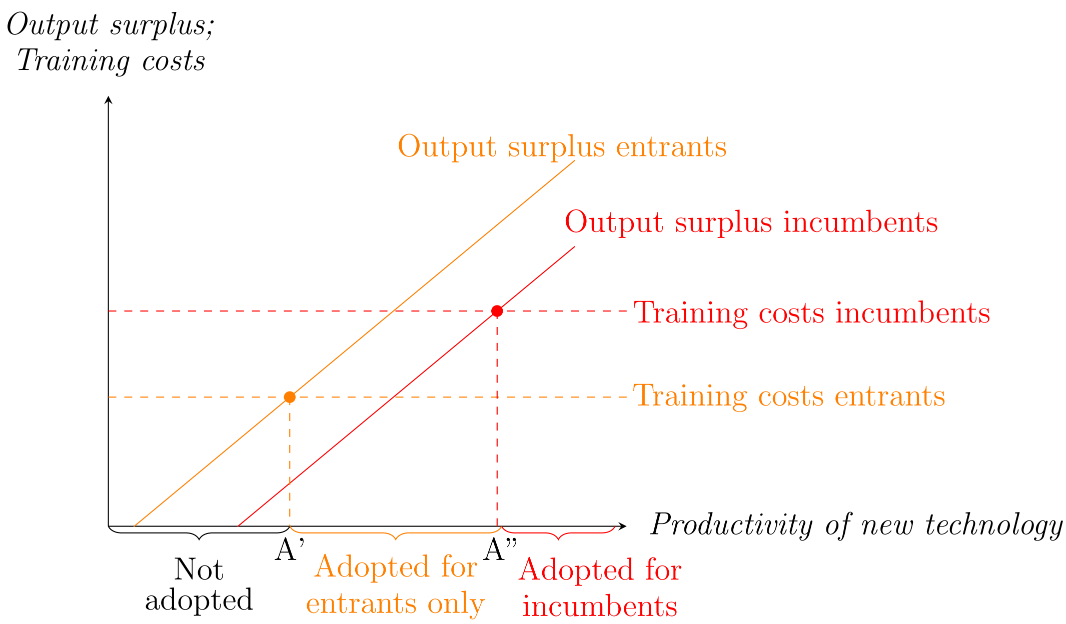
$$A_\tau \geq \left(1 + \frac{1}{1-\theta}\right) A_{\tau_0} \tag{D5}$$

Figure D1 visualizes this trade-off. New technologies below the productivity threshold A' are not adopted because training costs are too high, even for the least productive workers. New technologies above A' but below A'' are adopted by training labor market entrants only. New

[38] The renunciation of the assumption that wages are equal to marginal productivity is well backed up in the literature, in particular in the context of firm training (e.g. Konings & Vanormelingen, 2015).

technologies above the threshold A'' are adopted for incumbent workers as well.

Figure D1: Additional profits versus additional costs of technology adoption

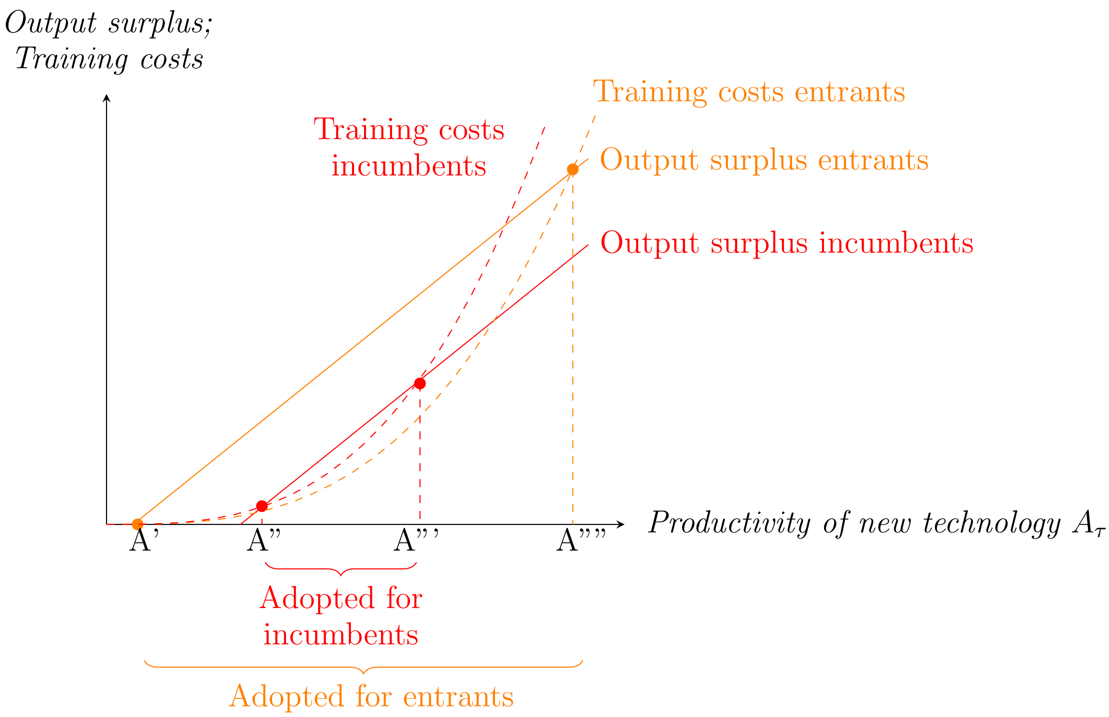


Notes: Profitability of training entrants versus incumbent workers.

Negative supply shock of trainees. Assume there is a missing entry cohort in $t = 1$. In this period, firms will invest in the new technology if and only if the new technology is productive enough to make it profitable to retrain incumbent workers. For productivity levels of the new technology $A' \leq A_\tau < A''$, this implies a reduction in firms' technology adoption compared to the case without a missing entry cohort.

Extension A – Increasing and convex capital adjustment costs. Equation (D2) implies constant capital adjustment costs of one period of training for any productivity level of the new technology. In standard capital adjustment costs models, adjustment costs are usually assumed to be increasing and convex in investment size, implying that small investments can be easily incorporated in the structure of the firm without much (re-)training, while big investments create larger disruptions requiring more training. Let us now assume that training costs are increasing and convex in technology productivity, $C'(A_\tau) > 0, C''(A_\tau) > 0$ in the sense that more productive technologies require longer periods of training. Due to the convex relationship, there are new technologies that require prohibitively long training, exceeding the productivity gains. The productivity level above which adoption is prohibitively costly is reached earlier for incumbent workers because their training is more expensive. The trade-off between additional profits and additional costs of production are shown in Figure D2. A new technology is adopted for entrants only below a certain productivity threshold A'' and above a certain productivity threshold A'''. In consequence, a lack of entrants hinders the adoption of little productive technologies, as in the setting with constant training costs, and the adoption of very productive technologies. If productive technologies are more costly, this implies larger investment losses if entrants are missing.

Figure D2: Additional profits versus additional costs with convex adjustment costs



Notes: Profitability of training entrants versus incumbent workers when capital adjustment costs of training are increasing and convex in technology productivity.

Extension B – Infinite time horizon and worker retention. I now allow firms and workers to live for an infinite time horizon and workers to switch firms. I abstract from worker retirement which would present an additional factor why training entrants is in expectation more profitable than retraining incumbents. For simplicity, assume there is no temporal discounting or capital depreciation. The expected total surplus of adopting a new technology τ is now given by the sum of all expected future output increases net wage increases, minus one-time capital adjustment costs and capital costs:

$$E[\Delta\pi_{j\tau}] = E[\Delta Y_{j\tau} - \Delta W_{j\tau}] - C_{j\tau} \tag{D6}$$

Workers can leave their firms at the end of each period. The probability of a worker to stay at a firm j, p_j, is inversely related to the firms' monopsony power and exogenously given.[39] For each worker group, the net outplus surplus from technology adoption extincts as soon as

[39] Firm monopsony power may include classic monopsony aspects such as concentration or outside options, but also aspects related to firm training, such as information asymmetries about worker skills. See the excellent survey by Wolter & Ryan (2011). For the purpose of this study, the underlying reasons are irrelevant, and I assume p_j to be exogenously given.

this worker group is retrained in a new technology. The retraining probability $\phi(p_j)$increases in the firms retention rate p_j, $\phi(p_j) > 0$. Hence, the expected total net output increase is equal to the net output surplus of technology adoption of each worker group in each (future) period multiplied by the probability of being at the firm and not being (re-)trained in this period:

$$E[\Delta Y_{j\tau} - \Delta W_{j\tau}] = (1-\theta) \sum_{t+1}^{T} p_j^t (1-\phi(p_j))^t \sum_{\tau_0=0}^{\mathcal{T}-1} \Psi_{\tau_0} L_{jt\tau_0} (A_\tau - A_{\tau_0}) \tag{D7}$$

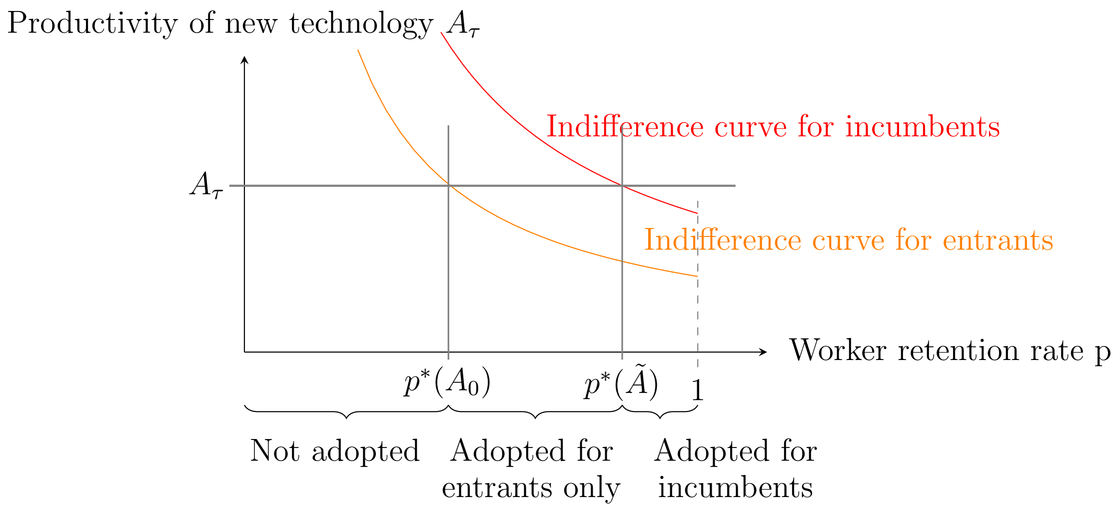
Technology adoption is more profitable the longer a firm benefits from trained workers, i.e. the higher p. Combining (D6) with (D2) and (D7), firms train a worker type L as long as the following condition holds:

$$A_\tau \geq \left(1 + \frac{1}{1-\theta} \frac{1}{\sum_t p_j^t (1-\phi(p_j))^t}\right) A_{\tau_0} \tag{D8}$$

Figure D3 visualizes the indifference curve along a firms' worker retention rate p for entrants with productivity A_0 (orange curve) and incumbents with productivity $\tilde{A}$ (red curve) for a given productivity level of a new technology at A_τ. This new technology is not adopted if the firms' retention rates is below the lower threshold $p^*(A_0)$. For retention rates above $p^*(A_0)$ but below the upper threshold $p^*(\tilde{A})$, the technology is adopted for entrants only. For retention rates above $p^*(\tilde{A})$, the technology is adopted for incumbents as well.

Similarly, we can fix the firms retention rate p and study the different productivity levels needed to make adoption by different worker groups profitable. For a given worker retention rate, adoption and training is more profitable the lower the initial productivity of the worker.

Figure D3: Indifference curves for Extension A



Notes: Firms' indifference curve between non-adoption and adoption of the new technology τ depending on firms worker retention rate p for two example levels of worker initial productivity A_{τ_0}.

Let us turn to the case when no entrants with A_0 are available. For a given technology, firms with a retention rate below $p^*(A_0)$ will not reduce their technology adoption compared to the counterfactual scenario with entrants, because they would not have adopted in the counterfactual scenario either. For firms with a retention rate above $p^*(A_0)$ but below $p^*(\tilde{A})$, technology adoption is lower than in the counterfactual scenario. For firms with a retention rate above $p^*(\tilde{A})$, technology adoption without entrants is still profitable and, hence, does not drop compared to the counterfactual scenario.

Missing entrants might be substituted with other untrained workers if workers from previous cohorts are still untrained because all past technologies were not productive enough, or if worker retention in some firms is below the lower threshold.