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

Cäcilia Lipowski*

[Link to latest version]

November 20, 2023

Abstract

Firms in developed countries increasingly report skill shortages. 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. The complementarity between trainees and technology adoption is explained by firms' low opportunity costs and high productivity gains of training young labor market entrants in skills required for technology adoption. These findings dampen hopes of counteracting labor shortages by substituting labor with capital.

Keywords: Labor Shortages, Firm Investments, Technology Adoption, Vintage-Specific Technical Skills

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

*ZEW Mannheim, caecilia.lipowski@zew.de

I am grateful for comments by Daron Acemoglu, Melanie Arntz, David Autor, Eduard Brüll, Katja Görlitz, Maarten Goos, Harald Pfeifer, Pascual Restrepo, Anna Salomons, Johannes Schmieder, Anna Waldman-Brown, Nicolas Ziebarth and Ulrich Zierahn-Weilage. 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 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, referring to reductions in labor supply, may have significant consequences for economic growth if, in response, firms adjust investments 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 shortages on firm investments is challenging because labor scarcity tends to evolve gradually; it usually goes hand in hand with changes in labor demand; and it is often confounded by unobserved factors at the region, industry, or firm level.

In this paper, I overcome this identification issue and provide empirical evidence on the causal effect of labor shortages on firm technology investments exploiting an education reform in Germany in 2001. I focus on 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. Indeed, in a representative survey, approximately half of the firms agree that in-house training of young labor market entrants ensures the firm's constant supply of new skills, improves the firm's innovative capacity, and enhances the firm's adaptability to technological changes.[1]

My identification strategy exploits a natural experiment created by an education reform. In 2001, two out of six East German federal states permanently increased the length of schooling required for the university entrance degree by one year.[2] In treated states, the reform caused a missing school graduation cohort that translated into a missing trainee entry cohort in the same year and a reduced the stock of trainees in subsequent years. The reform-induced variation in trainee supply across time and states is plausibly exogenous to firms, in particular since trainees are highly immobile.[3] The missing trainees can be thought of as not yet skilled but future middle-skilled professionals, often in white-collar occupations such as media, retail or financial service occupations.

The German vocational training system provides a unique opportunity for studying implications of shortages of young labor market entrants for three reasons. First, it is omnipresent in the labor market with two thirds of the workforce having a completed vocational training trainee. Second, its institutionalized set-up allows the precise identification of trainees, training firms and training periods in data. Third, trainee wages are rigid for institutional reasons, such

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

[2] Büttner & Thomsen (2015); Morin (2015); Muehlemann et al. (2018); 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.

[3] Only 2.2% of trainees move federal states for their vocational training (SOEP; own calculation). Likewise, only 5% of trainees commute between federal states (LIAB; own calculation).

that the reform creates an exogenous decrease in labor quantity, rather than an increase in the price of labor. At the same time, vocational training in Germany is comparable to on-the-job training in other countries, which is why I use the terms "young labor market entrants" and "trainees" interchangeably.

I compare investments and technological change of training firms in treated East German states to investments and technological change of similar training firms in untreated East German states by combining a difference-in-differences event study approach with firm-level matching. 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 as 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 has a substantial negative effect on firms' employment of trainees from the reformed school track, i.e. trainees with 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 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.

Second, investments decrease sharply in training firms in treated states compared to training firms in control states following the reform. While highly educated trainees represent only 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 the pivotal role trainees play in the adoption of technologies requiring new skills, as I clarify in the following, coupled with the highly right-skewed distribution of investments. In fact, the estimated average decrease comprises firms refraining from large investment projects and firms without investment reductions since they would not have invested in the counterfactual scenario. This finding is in line with the literature emphasizing the lumpy nature of investments. Importantly, the reform-induced investment reductions are not recouped in the following years, leading to a permanent reduction in the capital stock in treated training firms even though the supply shock is temporary in nature.

I confirm the link between the investment decline and the absence of trainees in several ways. I provide evidence that non-training firms in non-exposed industries do not reduce their investments. Also, I show that 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. Additionally, I exploit firm-level exposure to the reform as an instrument for trainee employment in a complementary identification strategy. Confirming the relationship between trainees and investments, I find that firms more affected by the negative trainee supply shock decrease investments to a greater extent.

The third key finding establishes that the investment decline is linked to reduced technology adoption. Due to the negative trainee supply shock – hitting in a time period of strong technological and skill change – the technical condition of machinery depreciates in treated training firms. 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). Investments in production technologies and information and communication technologies (ICT) become less likely.

To rationalize the effects of the negative supply shock of young labor market entrants on firms' technology investments, I present a stylized economic framework which can explain my findings under two conditions. First, technologies require up-to-date skills, and second, in expectation, trainees stay at their training firm long enough to redeem the firm's investment in their human capital. Under these assumptions, it is more profitable for firms to train young labor market entrants in tech skills than to retrain incumbent workers because opportunity costs of training young labor market entrants are lower and productivity gains are higher. 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, it may not be profitable for firms to adopt technologies requiring new skills.

I provide empirical evidence in support of both assumptions. If the trainee shortage impedes technology adoption due to the need of new, technology-specific skills, firms more exposed to skill changes should slow-down technology adoption more when trainees are scarce. Based on Lipowski et al. (2023), who demonstrate that skills taught during vocational training are updated to technological progress, I document that firms with higher pre-reform shares of workers with outdated skills indeed decrease investments more.

The second assumption is that workers stay at their training firm for a sufficient amount of time. Indeed, approximately 40% of trainees remain at their training firms upon training completion. Also, I show that investments drop more in firms with higher trainee retention rates, i.e. firms employing trainees as an investment in skills for future production.

This paper is the first to show that shortages of young labor market entrants causally and significantly decrease firm technology investments. I hereby contribute to three related literatures.

The most closely related strand of literature studies how technological change responds to economic incentives. Theoretical models highlight that technology invention and adoption is endogenous to the relative abundance of production factors (e.g. Zeira, 1998; Acemoglu, 1998, 2002). Empirical papers, mainly relying on migration shocks, confirm this theory. A decrease (increase) in the supply of low-skilled labor positively (negatively) affects labor-saving patenting (Dechezleprêtre et al., 2019; Danzer et al., 2020; Andersson et al., 2022; San, 2023) and fosters (reduces) the adoption of production technologies in agriculture and manufacturing (Lewis, 2011; Hornbeck & Naidu, 2014; Clemens et al., 2018). 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 provides empirical microevidence for endogenous technology adoption in the context of a negative supply shock of young natives. It benefits from a clear identification that is free from concomitant labor demand effects common to migration or fertility shocks, and migrant-specific skills. Also, I move beyond aggregate effects, showing that firms facing stronger shortages adjust investments more. The findings imply that young labor market entrants are complementary to technology adoption, suggesting that the effects of shortages of young workers on economic growth are likely more detrimental than previously surmised.

Second, I contribute to the literature on technology-specific human capital. The literature provides many examples of how new technologies require new skills, 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 technology-specific skills have been linked to decreasing returns to experience, early retirement, and reduced hiring opportunities for older workers (Aubert et al., 2006; Ahituv & Zeira, 2011; Deming & Noray, 2020). Adão et al. (2020) show that the adaptation to technologies that require technology-specific human capital takes place through the entry of young workers, rather than by upskilling incumbent workers. They show that this particularly applied to the ICT revolution studied in this paper. My paper is the first to incorporate technology-specific skills into the concept of endogenous technological change, highlighting that the scarcity of young labor market entrants hinders the adoption of technologies requiring technology-specific skills.

Third, I contribute to the nascent literature on the consequences of labor shortages on firm outcomes. Existing studies establish a negative effect on firm capital, sales, and productivity (e.g. D'Acunto et al., 2020; Le Barbanchon et al., 2023; Sauvagnat & Schivardi, 2023). I propose a novel mechanism through which reduced labor supply affects firm outcomes, namely the relevance of young labor market entrants for 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 investments (Section 6). Section 7 presents the economic rationale behind the investment drop and provides empirical evidence on the mechanism. 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). 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 calculation).[5] Among those firms 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.

[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. 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).

A representative firm survey suggests that trainees play a key role in firms' acquisition of new tech skills: when asked about their vocational training, 55% of the firms agree that it ensures the constant supply of new skills, 52% agree that it improves the firms' innovative capacity, and 47% 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. For identification, I exploit this labor supply reduction 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] While two thirds of the missing upper track school graduates would usually opt for university studies, I instead focus on those upper track school graduates who subsequently start vocational training. This endorses the credibility of the identification strategy because vocational trainees postpone their labor market entry less and move or commute less across federal states.

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.

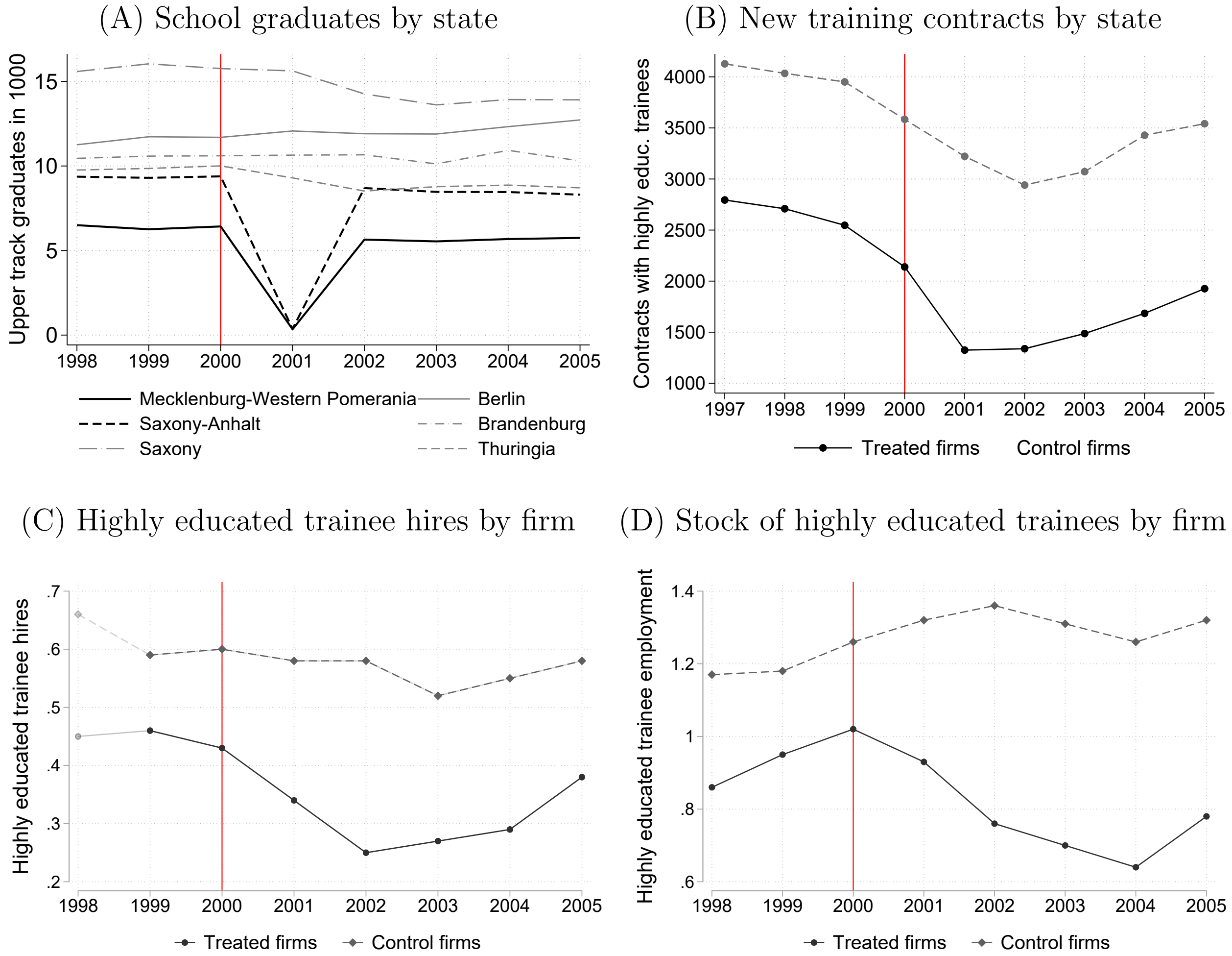
Based on the official training statistics, Panel B shows that the missing school graduation cohort translates into a missing cohort of newly concluded training contracts. Since at that time, males in Germany had to do military service of 10 months when reaching the age of 18,

[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.

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

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. Hirings in 1998 should be taken with caution. Own calculations.

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,

[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.

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 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 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

[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.

[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.

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 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 Table **??**. 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 2.8% of East German workers 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. Highly educated trainees are most common in the business service sector, but can also be found in the manufacturing or construction sector, see Appendix Figure A2.1.

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

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

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	.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. For further summary statistics see Appendix Table A2.1.

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^{***}
# highly educated trainees	.06	5.04	-4.98^{***}
% highly educated trainees	.06	2.45	-2.39^{***}
Inv. per worker (in €1,000)	14.61	18.80	-4.20^{***}
	Selected Industries		
Manufacturing	.33	.29	$.04^{**}$
Energy, water, waste	.03	.04	$-.02^{**}$
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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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 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, a training firm employed on average 5.04 highly educated trainees, corresponding to 2.45% of the firms workforce. Compared to non-training firms, training firms are three times as large in employment size, 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. Appendix Table A2.2 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.

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. In summary, capital investments occur regularly in the data.

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

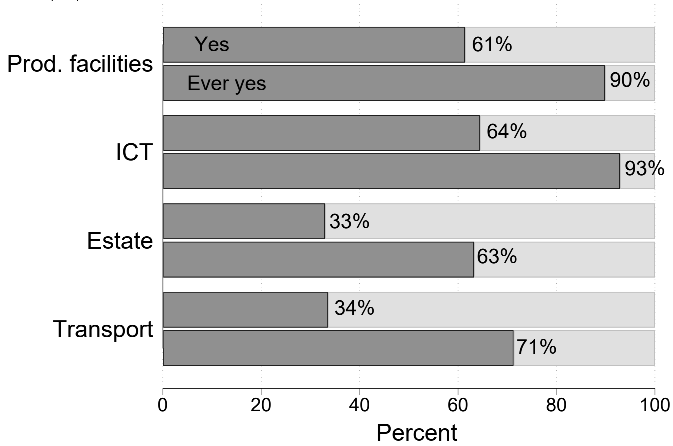
[16] Since inflation affects all firms equally, it will be absorbed by the year fixed effects.

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 A2.3. *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.

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.

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.

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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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:

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

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]

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 unmatched 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

[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.

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.

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

[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.

[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.

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 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.

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. Hence, confidence intervals based on standard errors clustered at the state level would not be valid. For valid inference, I follow two approaches suggested by Roth et al. (2023). First, 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 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 bite of the reform. The left panel shows the sample of unmatched firms, the right panel the sample of matched firms. Panel A shows a clear drop in highly educated trainee hires among 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

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

have appeared in the data as new hires.[21] With 0.64 fewer hires in 2002 in the sample of unmatched firms (-1.11 in the 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 fully 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 experienced 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 of 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. As expected, 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. 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 matched and the sample of unmatched training firms for both hires and stock of highly educated trainees.

Wage and substitution effects. The data allows to study firm adaptation behavior such as wage changes or the substitution of highly educated trainees with other worker types. To investigate such accompanying effects, I employ the corresponding difference-in-differences specification, distinguishing the pre-treatment period 1998–2000, the roll-out year 2001, the post-treatment period 2002–2004, and the phase-out year 2005. Results are given in Table 4.

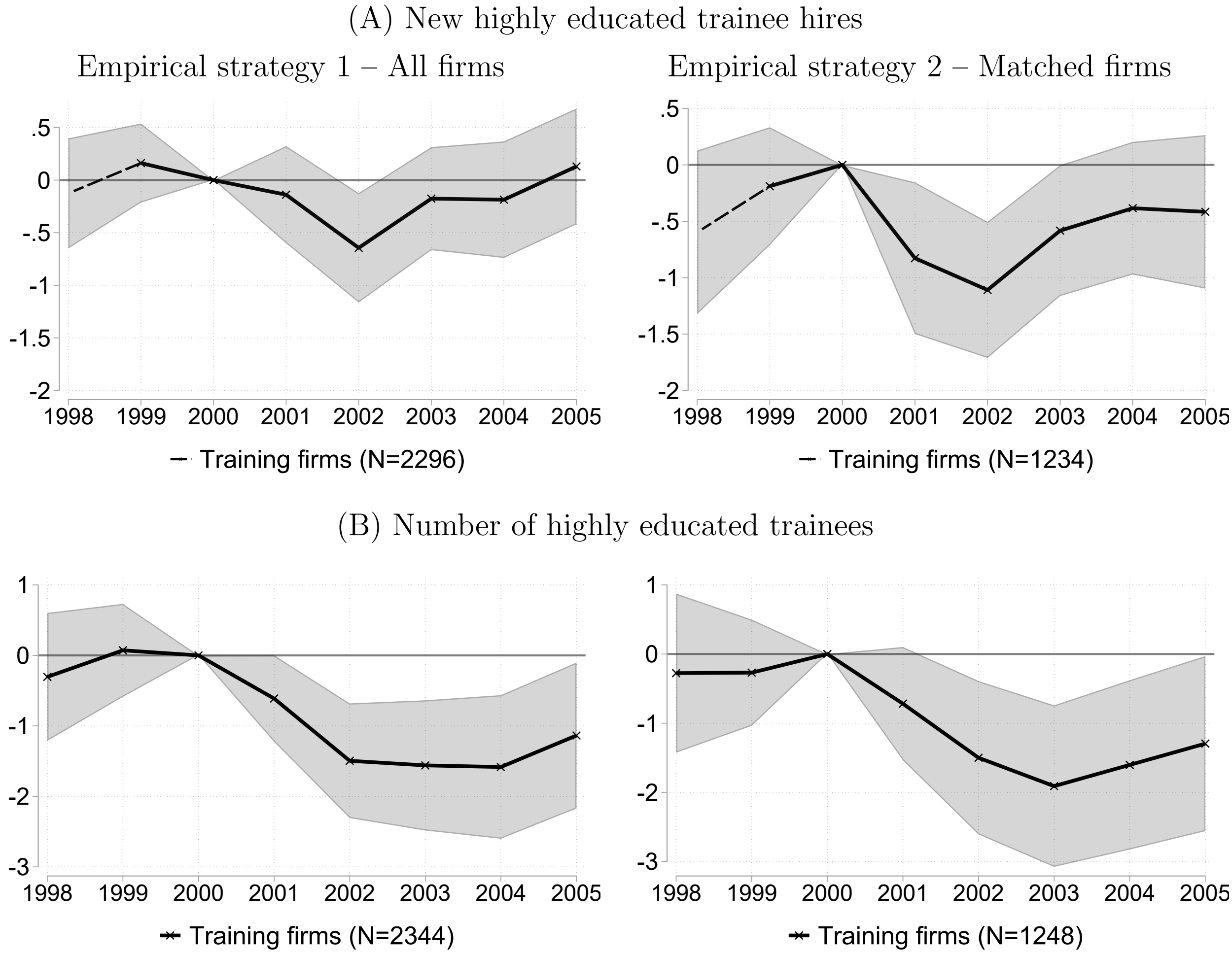
In contrast to what standard economic theory would predict, there is no evidence of an increase in wages of highly educated trainees in response to the negative supply shift (column 1). This result is in line with other papers emphasizing the stickiness of training wages (e.g. Muehlemann et al., 2018). Firms may shy away from increasing wages in response to a temporary shock because downward rigid wages will impede a subsequent wage decline as soon as the supply shock fades out. Also, in order to attract non-school graduates, trainee wages would need to increase by a very large amount, since trainee wages are only a small fraction of the wages of already qualified workers. Dorner & Görlitz (2020) argue that the estimated absence of a wage increase can be attributed to firm selection. The reason is that counterfactual wages

[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.

[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: Bite of the shock



Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Standard errors clustered at the firm level. For the corresponding difference-in-differences estimates see Table B2.1, Panel A, column 1 and 2. Hirings in 1998 should be taken with caution.

of firms that do not employ highly educated trainees due to the negative trainee supply shock are unknown. Indeed, accounting for selection, they find increases in training wages which are, however, very small in magnitude (1%). I follow their approach and control for firm selection by including firm fixed effects (column 2). The estimates remain virtually unchanged, indicating that wages of highly educated trainees did not increase also after controlling for firm selection even when controlling for firm selection.

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

Next, I turn to the number of highly educated trainee hires in treated versus control firms that is commuting from a different federal state. The coefficient of interest captures potentially increased commuting into treated states plus potentially reduced commuting into control states.

Table 4: DiD Results – Wage and substitution effects

	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)
Treated	*Empirical strategy 1 – All training firms*					
× Post	−0.03	−0.04	−0.39	0.03	−0.09*	−0.13**
	(0.04)	(0.03)	(0.82)	(0.05)	(0.05)	(0.05)
× Phase-out	−0.08*	−0.07*	0.68	−0.04	−0.14**	−0.16
	(0.04)	(0.04)	(1.04)	(0.10)	(0.06)	(0.10)
Firm FE		X				
Observations	1758	1758	2296	2018	2227	2344
Init. outcome	3.00	3.00	6.30	0.04	0.42	4.87
Treated	*Empirical strategy 2 – Matched training firms*					
× Post	0.01	−0.01	0.02	0.04	−0.07	−0.07
	(0.05)	(0.04)	(1.10)	(0.06)	(0.07)	(0.07)
× Phase-out	−0.05	−0.04	0.01	0.08	−0.12	−0.03
	(0.05)	(0.04)	(1.50)	(0.11)	(0.11)	(0.14)
Firm FE		X				
Observations	908	908	1234	1082	1190	1248
Init. outcome	3.00	3.00	5.97		0.43	4.82

Notes: Reference group: Treated × Pre. Pre: 1998-2000. Roll-out: 2001. Post: 2002-2004. Phase-out: 2005. Controlling for periods 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. For the full set of results, see Appendix Table B2.2. For further outcomes see Appendix Table B2.3.

There is no evidence of increased cross-state commuting of highly educated trainees following the shock (column 4), validating the SUTVA assumption.

Firms may also increase retraining of incumbent workers to overcome skill shortages. In contrast, I observe a decline in internal training measures in treated training firms by approximately one third (column 5). This finding might be related to the foregone introduction of new technologies which necessitate training, as I show below. Likewise, column 6 shows that there is no evidence that firms substitute trainees with workers with completed vocational training. 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 B2.3.

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

6 Effects on firm investments

6.1 Main effects

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 strongest investment decline is observed in the years 2002 and 2003, with €5,250 and €6,370 less per worker in the sample of unmatched firms. This corresponds to a drop of 30% and 37% of the 1998 value. To compare this value with other estimates in the literature, I turn to the effect on log investments for the sample of firms with strictly positive investments (see Appendix Table B2.4, column 5). The estimated decrease by 33 log points corresponds to the decline the literature would predict if capital costs increased by 9-15% (Zwick & Mahon, 2017; Lerche, 2019; Liu & Mao, 2019).[24]

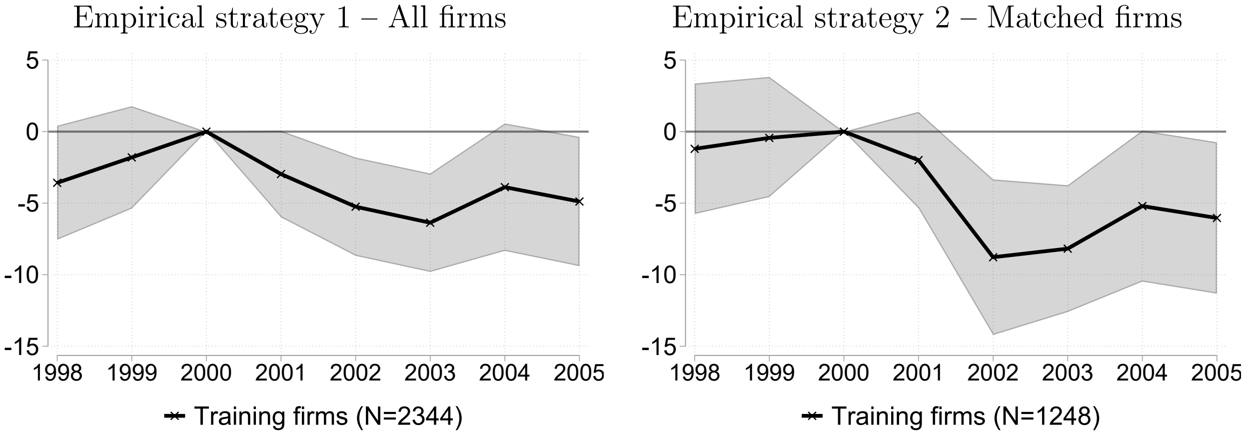
The negative effect diminishes after 2003, corroborating its relation to the temporary drop in trainee supply. However, investments do not fully return to their initial level by 2005, potentially because affected firms are placed on less favorable long-term trajectories. Since investments of treated firms do not surpass investments of control firms at least until 2005, the capital *stock* of treated training firms is likely permanently reduced despite the trainee supply shock being only temporary. There are no statistically significant pre-trends for the sample of unmatched firms. Within the sample of matched firms, the investment drop is even larger. To ensure that the effect is not driven by changes in firm employment, I turn to total (log) firm investments, see Appendix Table B2.4, columns 4 and 5. The corresponding coefficients are also negative, statistically significant and large in magnitude. 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 employment.[25] In

[24] 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.

[25] 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

Figure 4: Investments per worker in €1,000



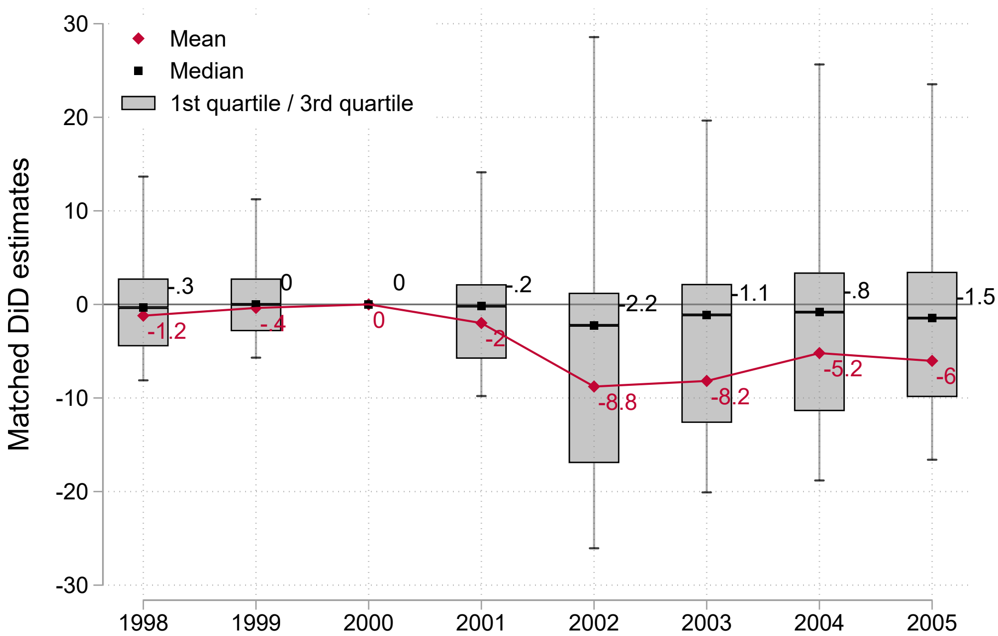
Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Standard errors clustered at the firm level. Outcome: investments divided by total employment in 1998 in €1,000. For the corresponding difference-in-differences estimate see Table B2.4.

fact, the average investment decline is remarkably large given that highly educated trainees present on average only 2.5% of a training firms' 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. Just as the distribution of investments, the distribution of the treatment effect on investment is highly right-skewed. The *average* investment drop in 2002 – the statistic reported throughout the paper – is four times larger (€-8,800) than the *median* investment drop (€-2,200, 12% of the 1998 value). This finding is in line with the literature highlighting the lumpy nature of investments (e.g. Cooper et al., 1999; Bessen et al., 2020): in the counterfactual scenario, some firms were not planning to invest, 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. The corresponding figure for total investments (Appendix Figure B2.1) shows a similar pattern.

To explicitly analyze the effect on large investments, I turn to the intensive and extensive margin of adjustment, see Figure 6. Regarding the intensive margin, I conduct an event study regression among observations with strictly positive investments using a binary outcome taking the value of one for large investments in the upper tercile of the overall investment per worker distribution (>€10,000) and zero otherwise. Treated training firms are 15 percentage points (24 percentage points in the matched sample, respectively) less likely to make large investments than control training firms in. 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 8. When defining large investments within industries, the effect remains, see Appendix Table B2.4, column 9. The overall effect is not purely driven by adjustments at the intensive margin.

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.

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

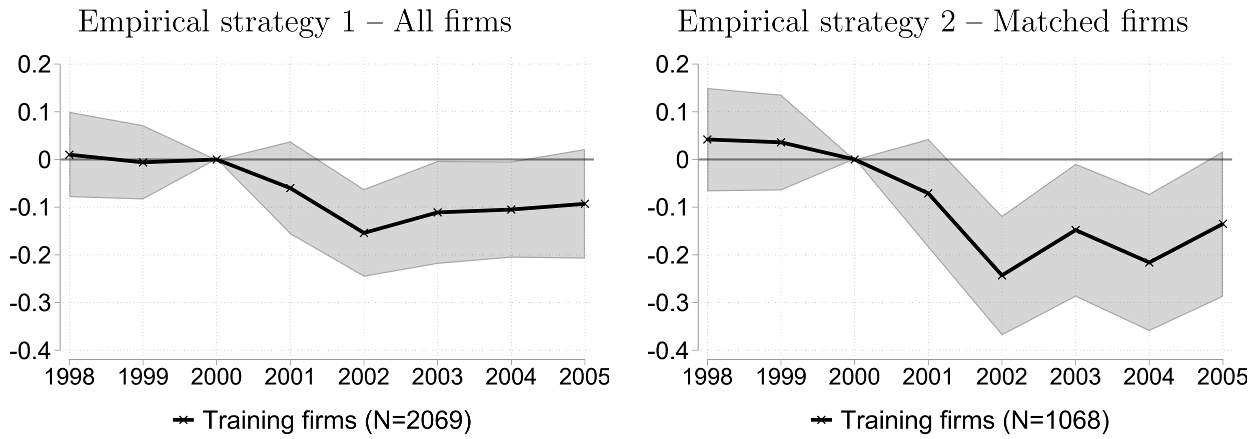


Notes: Distribution of the matched firm-level difference-in-differences estimates. Outcome: Investments per worker in 1,000. *Red:* Average. *Black:* Median. *Box:* 25th and 75th percentile. *Adjacent values:* 25th percentile - 1.5*the interquartile range (75th percentile + 1.5*the interquartile range).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

Treated training firms are also 5–12 percentage points more likely to not invest at all in 2003, see column (6) in Appendix Table B2.4.

Falsification test and industry spill-overs 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 directly affected by the reform. Confirming this hypothesis, the average investment drop among non-training firms is less than half of the average drop among training firms and statistically less significantly different from zero but not zero, see Table 5, columns 1 and 2. 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 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 or knowledge spill-over. To test this, I perform a difference-in-differences regression including the triple interaction term between Treated, Post, and industry exposure to the reform, i.e. the share of highly educated trainees in an industry, controlling for all corresponding two- and one-way interaction terms (columns 3, 5 and 7). This analysis reveals that there is no negative effect on investments among non-training firms in unexposed industries. Non-training firms in exposed industries, in contrast, also reduce investments.

Figure 6: Intensive investment margin: Large investment (1/0)



Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Standard errors clustered at the firm level. *Panel (A):* Outcome: large investment in the uppest investment tercile versus small but strictly positive investment below the uppest tercile (1/0). For the corresponding difference-in-differences estimate see Table B2.4, columns 6 and 7.

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 7, comparing the post-reform years 2002–2004 with the pre-reform years 1998–2000. I show the estimates for both the unmatched 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 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.

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

Table 5: Falsification test and industry spillover

	Training firms	Non-training firms					
		Def. 1		Def. 2		Def. 3	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Unmatched						
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 (0–100)			−2.34		−3.15		−6.45
			(3.96)		(4.18)		(7.92)
N	2344	8744	8744	8024	8024	2816	2816
	Matched						
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 (0–100)			−2.76		−4.14		−6.48
			(5.57)		(4.55)		(8.24)
N	1248	6112	6112	5616	5616	1824	1824

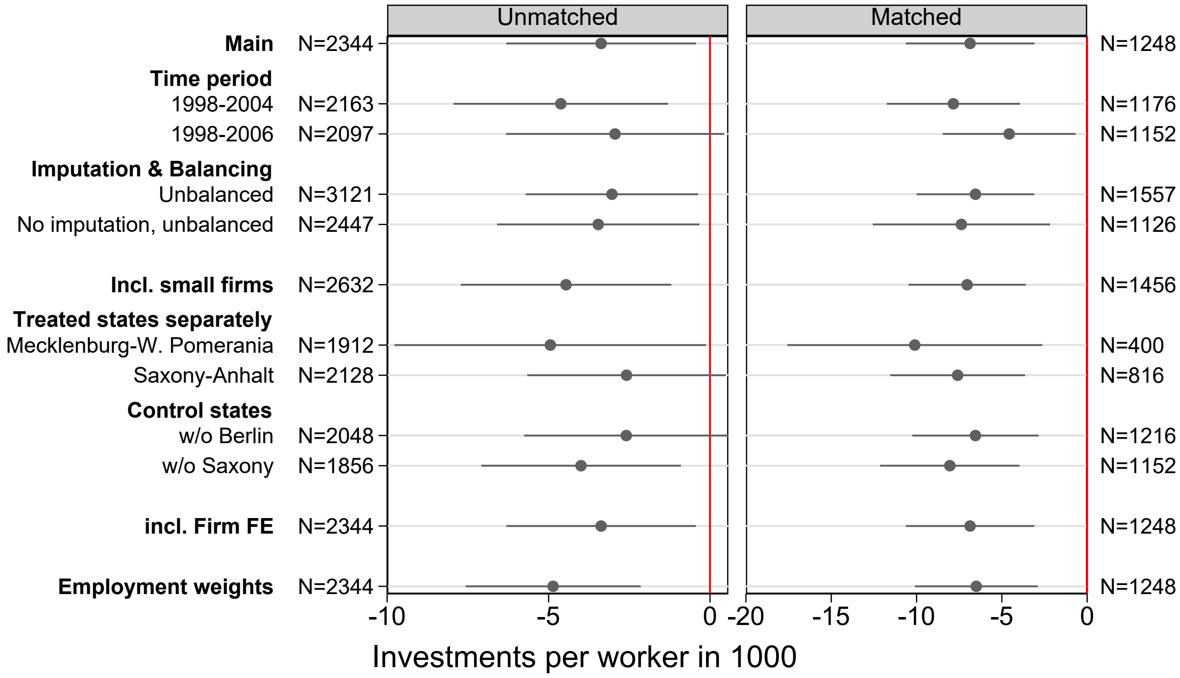
Notes: Outcome: investments per worker in €1,000. Industry exposure: Share of highly educated trainees at the detailed industry level in %. *Def. 1:* Firms without any highly educated trainee in 1998. *Def. 2:* Firms without any highly educated trainee in 1998-2000. *Def. 3:* Firms without any educated trainee in 1998-2000 independent of the schooling level. Reference group: Treated × Pre. Pre: 1998-2000. Roll-out: 2001. Post: 2002-2004. Phase-out: 2005. Controlling for periods and state fixed effects. Standard errors clustered at the firm level. See Appendix Table B2.4 for the full differences-in-differences estimation results of non-training firms and Appendix Figure B2.2 for the event study results.

small (e.g. Roth et al., 2023). Figure 8 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 cuased the observe investment decline. This result hold for both the matched and unmatched sample of 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 B2.5. 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.

Firm-level treatment intensity – Instrumental variable regression. The previous analyses distinguish between training and non-training firms. In addition, the data allow analyzing whether more exposed training firms, i.e. training firms that suffer from larger reform-induced trainee employment decreases, reduce investments more. As a complementary analysis,

Figure 7: Robustness - Investments per worker in €1,000

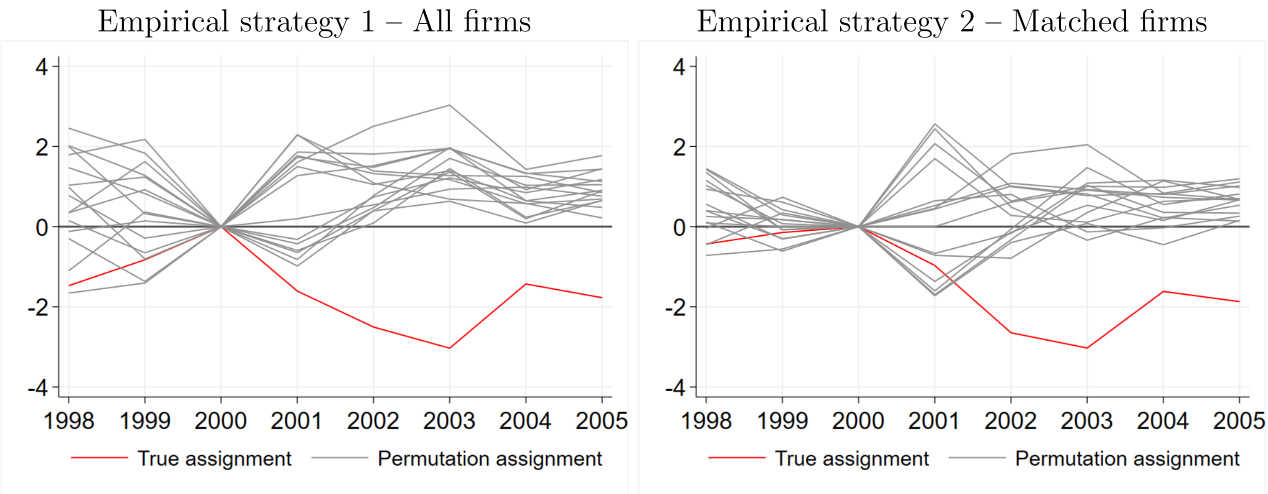


Notes: DiD estimates plus 90% confidence bands of the term Treated × Post compared to Treated × Pre plus number of observations (N). Post-reform: 2002-2004. Pre-reform: 1998-2000. Standard errors clustered at the firm level. Training firms only. *Main:* Main specification. *Time period:* requiring a 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.

I therefore perform an instrumental variable regression which predicts firm-level supply of highly educated trainees based on firms' initial employment (i.e. exposure) of highly educated trainees. 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 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, again, suggests spill-over effects among firms with no or few highly educated trainees within treated states.

Figure 8: 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 B2.4.

6.2 Effects on firm technology adoption

So far, I have extensively shown that a negative trainee supply shock importantly reduces investments. I next relate this investment drop to forgone technological change. I use the same difference-in-differences event study design as for the effect on firm investments. Results are given in Table 6.

As a direct 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 unmatched sample 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 B2.5.

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 6: 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)
Treated		*Empirical strategy 1 – All training firms*				
× 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)
× Phase-out		−0.27	−0.11*	−0.14**	−0.05	−0.07
		(0.17)	(0.06)	(0.06)	(0.05)	(0.06)
Observations	2341	1311	2341	2344	2344	2344
Init. outcome	3.97	1.35	0.72	0.80	0.59	0.35
Treated		*Empirical strategy 2 – Matched training firms*				
× 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)
× Phase-out		−0.20	−0.06	−0.03	−0.01	−0.05
		(0.24)	(0.09)	(0.07)	(0.08)	(0.10)
Observations	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 periods 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 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 unmatched sample, 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 B2.5.

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 Framework and supporting evidence

7.1 Framework

To rationalize the negative effect of the shortage of young labor market entrants on firm technology investments, I outline an economic argument that emphasizes the role of yet untrained labor market entrants in learning skills required for the adoption of new technologies. A more detailed formalization of the framework is available in Appendix D.

Upon arrival of a new, productivity-enhancing technology vintage, firms decide whether to adopt the new vintage or not in order to maximize profits. I assume that technologies require skills specific to each vintage, such that only those workers benefit from the technology-induced productivity increase that possess the corresponding skills. Note that this does not preclude technologies from simultaneously automating existent tasks.

Firms can acquire skills either by retraining incumbent workers or by training young labor market entrants within a vocational training program.[26] 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 retraining incumbent workers because their opportunity costs of training in terms of foregone production are lower and their productivity increase due to training higher.

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. In line with this, I find that firms with a 10 percentage point higher retention rate invest €400 (2.6%) more per worker and year, and have a significantly more up-to-date technical status of machinery, see Appendix Table B3.1.[27]

When young labor market entrants become temporarily unavailable, firms adopt a new technology only if the productivity gain is large enough to outweigh the costs of retraining incumbent workers. If retraining incumbent workers is too costly compared to its payoff,

[26] 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, such as having to invest in firm-specific skills, high hiring costs, or increased risk of bad personnel decisions/less opportunities for screening.

[27] By contrast, it is hard to test whether firms engaged in vocational training invest more than firms not engaged in vocational training because firms not employing vocational trainees may still train young university graduates.

technologies which would have been adopted if trainees were present, are not adopted. This aligns with the empirical finding that internal retraining of incumbent workers does not increase following the reform.

There are two alternative explanations for the complementarity between young labor market entrants and technology adoption beyond 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). Second, young workers tend to possess more recent skills (e.g. MacDonald & Weisbach, 2004). While I cannot rule out these two channels, they are unlikely to fully cause the observed investment decline. The reason is the sharp timing of the shock, with only marginally older workers from the previous training cohort still being available. If one of these two alternatives was at play, firms could simply retrain trainees of the previous cohort, since they are are only marginally less equipped with recent skills, and only marginally closer to retirement. The only aspect new labor market entrants are considerably different to second-year trainees is in their opportunity costs and expected payoff of training. This is supported by the Cost-Benefit Surveys of Vocational Training which show that firm revenues from skilled labor activities of second-year trainees are 134% higher than of first-year trainees, and 254% higher for third-year trainees than for 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 following, I provide empirical evidence in support of both assumptions, and therefore in support of this mechanism.

7.2 Empirical evidence regarding the mechanism

New skills. The literature provides many examples of how new technologies require new skills, without ruling out the replacement of existent tasks (e.g. Autor et al., 2003; Acemoglu & Restrepo, 2018; Deming & Noray, 2020; Autor et al., 2022). In the same setting as this paper, Lipowski et al. (2023) show that in Germany during the years around the reform, vocational skills as laid down in training curricula indeed changed frequently and due to technological change. However, not all occupations, and hence not all firms, are equally affected by technology-related skill changes. 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 relying on 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. (2023). This measure is ideal because it applies to the studied worker group, i.e. trainees, and because training curricula 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 whose training curricula are updated in 2000 and 2001. This concerns 18 occupations. 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.

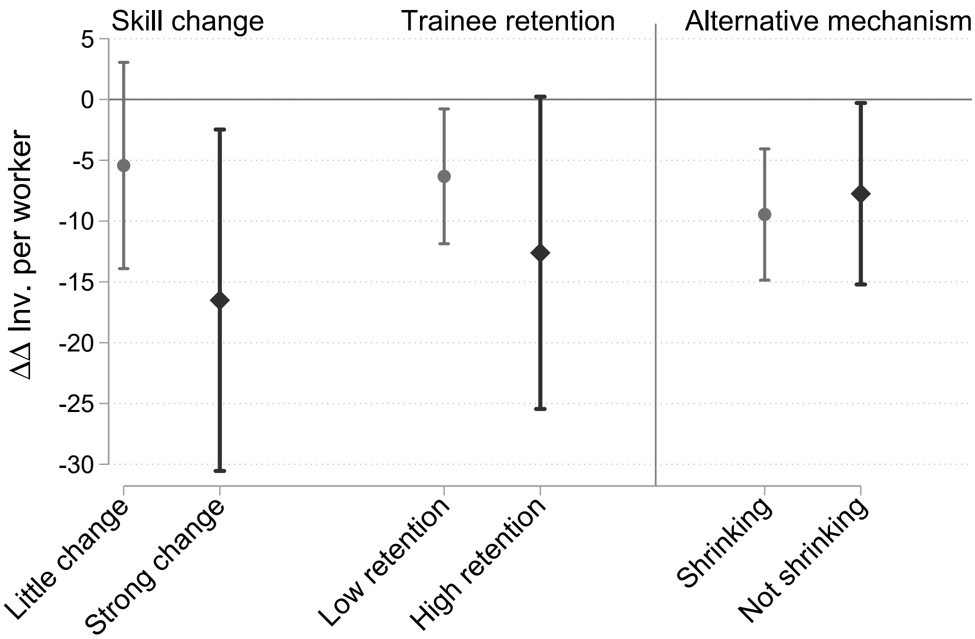
$$\Delta\Delta\text{Inv}_j = \alpha\text{NewSkills}_j + \beta X_{jt} + u_j \tag{3}$$

This approach is equivalent to a triple difference-in-differences specification with the triple interaction term Treated $\times$ Post $\times$ 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 for 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.[28]

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) is shown in Figure 9. As expected, firms with little skill change do not decrease investments due to the reform-induced trainee shortage. By contrast, investment declines are large for firms with strong skill change. This finding is robust across different specifications of exposure to new skills, see Appendix Table B3.2. These findings also suggest that the German vocational training system with its mandatory curricula and its skill transfer via vocational schools is an important driver of firm technology adoption. In line with this, Schultheiss & Backes-Gellner (2022) show for Switzerland, a country with a vocational training system similar to Germany, that changes in training curricula result in firms being closer to the technology frontier.

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

Figure 9: 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 my dataset is high with on average approximately 40% of the trainees remaining at their training firms.[29] However, there is variation in the retention rate across firms, see Appendix Figure B3.1, providing an opportunity to test this hypothesis empirically. I follow Mohrenweiser & Backes-Gellner (2010) and use a firm's trainee retention rate to draw conclusions regarding the firm's training strategy.[30] 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 the reform-induced investment reduction is indeed 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 outlined in equation (3), but with the trainee retention rate as the

[29]LIAB, own calculations. In general, German job tenure rates are comparable to the average rates in the European Union, but higher than the OECD average and the rate in the U.S. (OECD, 2000).

[30]Mohrenweiser & Backes-Gellner (2010) show that the distribution of the trainee retention rate in Germany is bimodal, with 44% of the firms retaining all trainees and 19% not retaining any trainee upon training completion.

independent variable of interest. The retention rate is defined as the proportion of apprentices staying at the firm in relation to all of the firms' apprenticeship graduates.[31] Figure 9, Panel B, shows the predicted investment changes for firms with high and low trainee retention rates. As expected, firms with high retention rates reduce investments heavily in response to the reform, while treated firms with low retention rates do not reduce their investments. For the regression table including robustness checks regarding the definition of the trainee retention rate, see Appendix Table B3.3.[32]

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 alternative mechanisms. Alternatively, a shortage of young labor market entrants may decrease firm investments because it creates an impediment to firm growth. If this was the primary driver behind investment cuts, only treated firms that do not substitute the missing trainees, i.e. firms that indeed experience a net reduction in their workforce, should reduce investments. In contrast, treated firms that replace the missing trainees with workers of other types should not reduce investments. I define not substituting firms as firms with a zero or negative absolute employment growth between 2000 and 2002, and as substituting otherwise. Figure 9, shows that investments decline similarly strongly for substituting and not substituting firms. For the regression table, see Appendix Table B3.4. This finding is incompatible with the alternative mechanism via barriers to 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 investments.

8 Conclusion and discussion

In this paper, I empirically demonstrate that a negative supply shock of young labor market entrants, especially those undergoing training, significantly reduces firm investments. This decline in investment is closely tied to a reduced adoption rate of technologies that demand up-to-date skills. My findings imply complementarity between young labor market entrants and technology adoption due to low opportunity costs and high expected pay-offs of young labor market entrants in acquiring new skills. Assuming that even labor-replacing technologies require some up-to-date skills, these results cast doubt on the feasibility of mitigating labor

[31] 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.

[32] Since the economic framework, see Appendix D, predicts a U-shaped pattern of investment decline by 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.

shortages through capital deepening (e.g. Acemoglu & Restrepo, 2018). They also 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).

While the natural experiment leveraged in this paper offers a unique strategy to identify the causal effect and explore its underlying mechanism, it may challenge the external validity of the results. In particular, the shock under examination is temporary, whereas the leading cause of current labor shortages, demographic change, is longer lasting. In general, all investments in times of scarce labor market entrants, no matter whether temporary or permanent, require retraining of incumbent workers that are associated with substantial additional costs. According to my economic framework, even when expecting a long-lasting labor shortage, my findings mimic firm behavior in the first year of labor shortage. With every year the shortage lasts, the reduction in firm investments become smaller because payoffs of technology adoption increase.

External validity also depends on the context in terms of the economic conditions. The labor supply shock studied in this paper hit in a period of a high unemployment rate of 18.8% (Federal Statistical Office, 2022) and an excess trainee supply (Ministry of Education & Research, 2004). The estimated effect therefore likely presents a lower bound of the effect a comparable shock would have in a tighter labor market.

Last, the investment decline might be amplified by factors specific to the German vocational training system. The definition of training curricula at the national level and the accompanying vocational courses in school enhance the transfer of new skills to trainees. Since the negative effect of trainee shortages on technology adoption is driven by the need for new skills, these aspects may result in a more pronounced effect compared to countries that do not actively promote the acquisition of new skills by trainees. Put differently, the results suggest that the German vocational training system serves as an effective catalyst for fostering the adoption of new technologies.

From a policy perspective, my findings not only stress the importance of expanding measures to attract and mobilize young labor market entrants. They also call for subsidies for retraining experienced workers when young labor market entrants are scarce, and for alternative channels to ensure skill transfer. 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.

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., 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.

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.

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

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.

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.

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.

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. (2023). 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., Pfann, G. A., Pfeifer, H., & Dietrich, H. W. (2018). The effects of supply shocks in the market for apprenticeships: Evidence from a german high school reform. Available at SSRN 3111138.

OECD. (2000). Employment by job tenure intervals - persons. Labor Force Survey. (14/08/2023)

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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

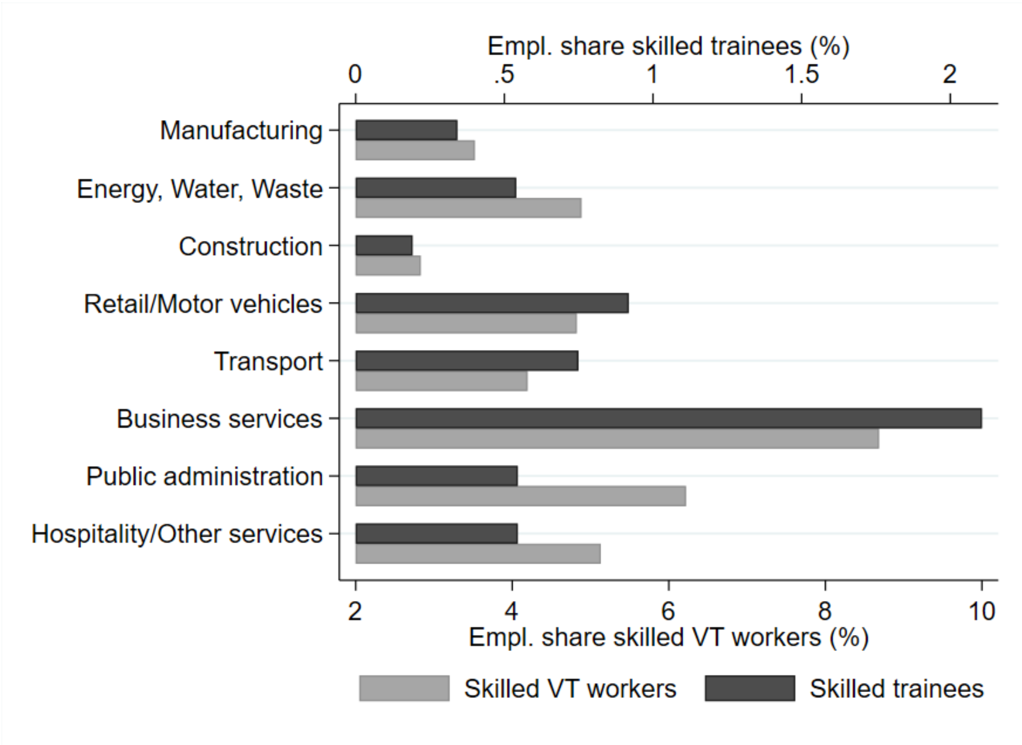
A.2 Descriptives and summary statistics

Table A2.1: Summary statistics

	Overall				Treated 1998		Control 1998	
	Mean	SD	Min	Max	Mean	SD	Mean	SD
	Employment & Wages							
# workers	148	325	10	9,570	153.5	221.3	176.5	420.2
# highly educated trainees	1.12	4.1	0	60	.86	3.19	1.17	4.42
# highly educated trainee hires	.47	1.95	0	36	.3	1.5	.51	2.49
Share highly educated trainees	.61	1.96	0	41.67	.47	1.81	.61	1.94
Overall wage	64.86	19.6	8.9	148.7	58.15	15.87	61.45	17.4
Wage trainees	18.58	6.02	0	92.75	17.54	6.46	17.64	5.66
	Industry							
Agriculture	.05	.22	0	1	.06	.23	.05	.21
Manufacturing	.32	.47	0	1	.25	.44	.35	.48
Construction	.1	.29	0	1	.1	.31	.09	.29
Retail/Motor vehicles	.09	.29	0	1	.1	.31	.08	.28
Business services	.13	.33	0	1	.13	.34	.13	.33
Public administration	.18	.38	0	1	.2	.4	.17	.37
Other services	.07	.26	0	1	.07	.26	.07	.26
	Investments							
Inv. per worker (in €1,000)	15.71	75.09	0	3932.86	16.88	38.25	20.09	57.64
Prob. to invest	.81	.39	0	1	.84	.37	.89	.31
Total investments (in €1,000)	2,945	16,948	0	888,827	3,242	9,839	3471	10,652
Inv. in prod. facilities	.61	.49	0	1	.63	.48	.69	.46
Inv. in ICT	.64	.48	0	1	.66	.47	.69	.46
Inv. in real estate	.33	.47	0	1	.39	.49	.4	.49
Inv. in transport	.34	.47	0	1	.38	.48	.38	.49
Tech. state of machinery	3.8	.75	1	5	3.82	.86	3.88	.76
Organizational changes	.67	.97	0	4	1	1.19	1.02	1.13

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

Figure A2.1: Highly educated trainees by industry

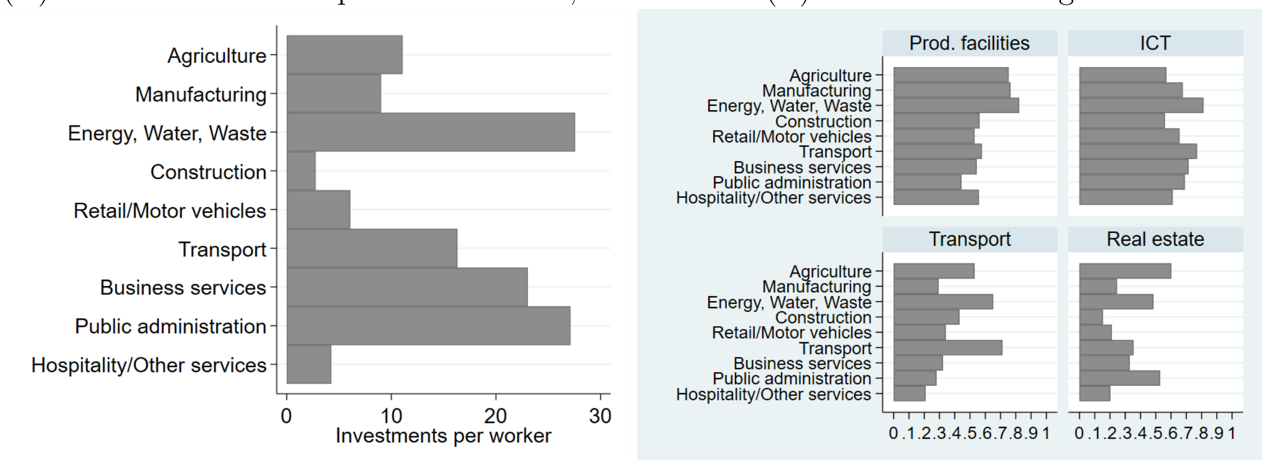


Notes: Share of highly educated trainees (highly educated VT workers) among total firm employment. Observations at the firm-year level. VT=Vocational training.

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 A2.2: Investment and technology indicators in the establishment panel

Variable	Survey Question	Manipulation	Frequency
Inv. per worker	What was the approximate sum of all investments [...]?	Divided by number of workers in 1998 from the administrative records and excluding outliers	Yearly
Inv. in ICT	Did your establishment invest in one or more of the following areas in the last business year? 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? 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? 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? 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 A2.3: Sum of total investments

Share of obs. w/o investments	19.3%
5th percentile	€10,000
25th percentile	€61,224
50th percentile	€331,633
75th percentile	€2,200,000
95th 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

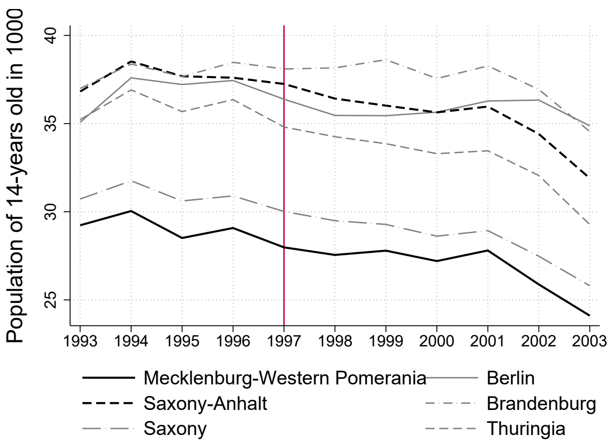
Figure B1.1: Employment of workers with completed vocational training/university studies

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

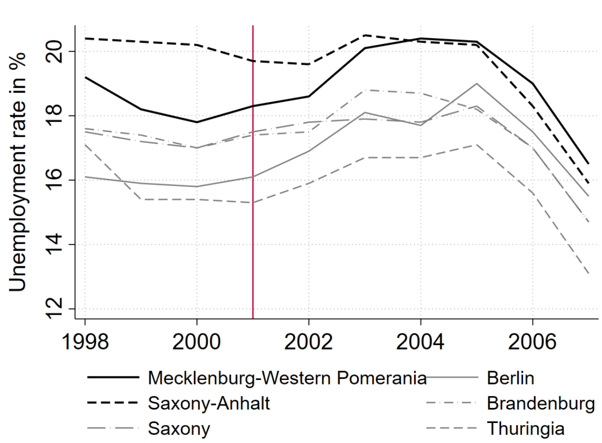
Notes: LIAB, own calculations. VT – Vocational training, TE – tertiary education. Red vertical solid line: Last year before the reform-induced trainee supply shock. Red vertical dashed line in Panel A: Last year before the reform-induced shock of workers with completed vocational training arrives (conditional 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

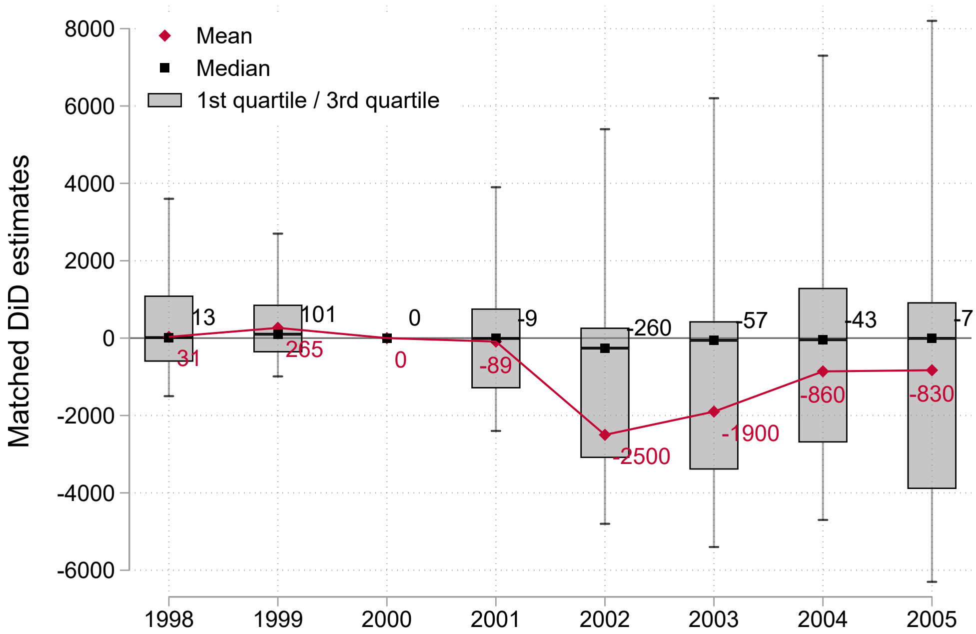
Table B1.2: Imbalance before and after matching – Untargeted variables

	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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

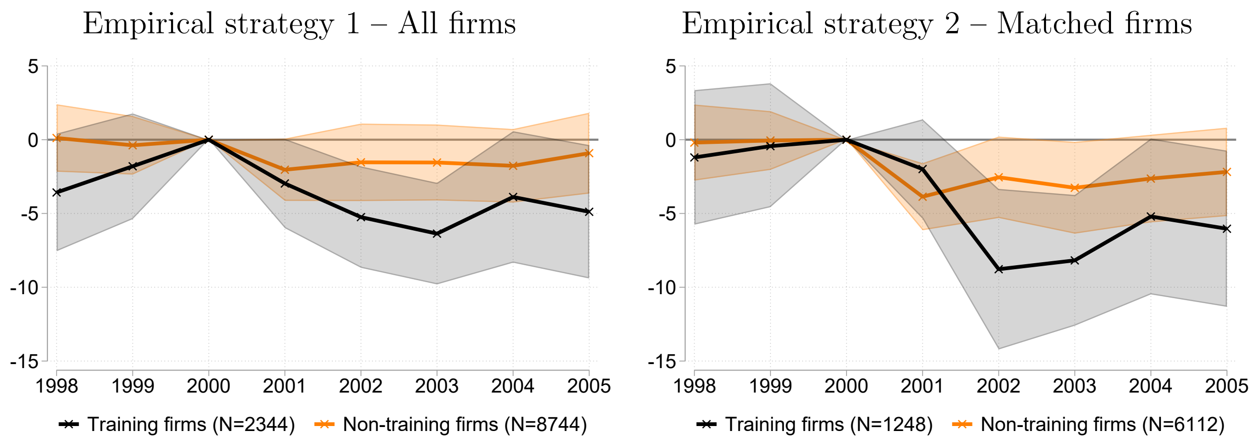
B.2 Additional results – Estimation results

Figure B2.1: Distribution of matched DiD estimates – Total investments



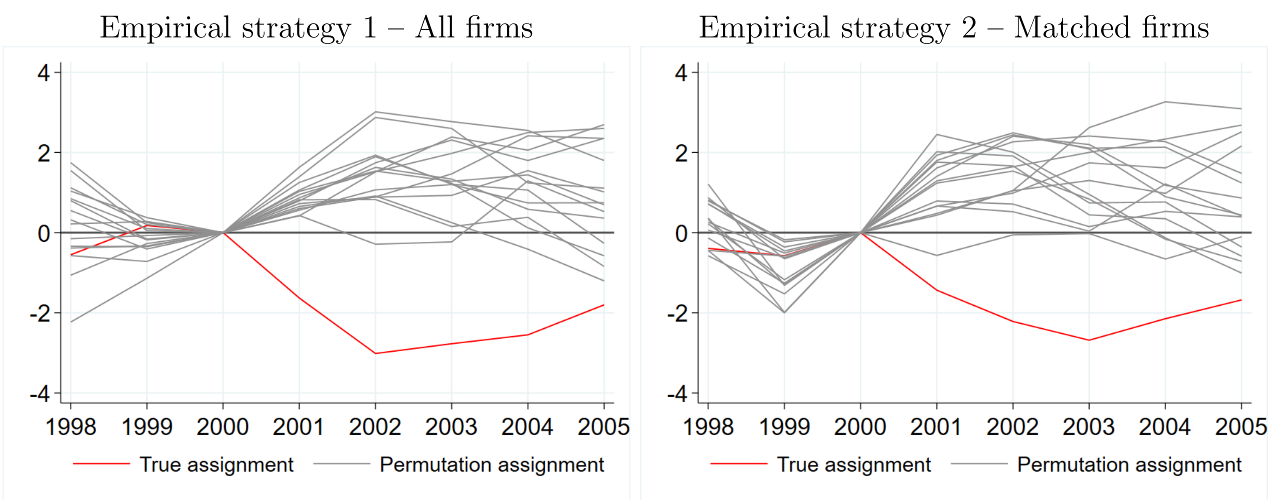
Notes: Distribution of the matched firm-level difference-in-differences in investments per worker. *Red:* Average. *Black:* Median. *Box:* 25th and 75th percentile. *Adjacent values:* 25th percentile-1.5*the interquartile range (75th percentile +1.5*the interquartile range). For the corresponding difference-in-differences estimate see Table B2.4, column 4.

Figure B2.2: 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 B2.4.

Figure B2.3: 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 8.

Table B2.1: DiD Results - Bite of the shock

	# highly educated trainees	# new highly educated trainee hires	% of highly educated trainees in total employment
	(1)	(2)	(3)
Treated	*Empirical strategy 1 – All training firms*		
× Roll-out	-0.53	-0.15	0.23
	(0.39)	(0.28)	(0.26)
× Post	-1.47***	-0.35	-0.31
	(0.49)	(0.27)	(0.27)
× Phase-out	-1.06*	0.12	-0.14
	(0.58)	(0.30)	(0.28)
Observations	2344	2296	2344
R^2	0.03	0.03	0.02
Init. outcome	4.9	2.24	2.69
Treated	*Empirical strategy 1 – All non-training firms*		
× Roll-out	-0.05	-0.06**	-0.14**
	(0.04)	(0.03)	(0.06)
× Post	-0.14***	-0.08***	-0.18***
	(0.04)	(0.02)	(0.05)
× Phase-out	-0.18***	-0.08**	-0.19***
	(0.06)	(0.03)	(0.07)
Observations	8744	8382	8744
R^2	0.02	0.01	0.02
Init. outcome	0.00	0.04	0.00
Treated	*Empirical strategy 2 – Matched training firms*		
× Roll-out	-0.54	-0.56	-0.08
	(0.51)	(0.38)	(0.32)
× Post	-1.49**	-0.43	-0.61*
	(0.68)	(0.29)	(0.03)
× Phase-out	-1.11	-0.15	-0.48
	(0.75)	(0.37)	(0.30)
Observations	1248	1243	1248
R^2	0.04	0.04	0.03
Init. outcome	4.90	2.26	2.76
Treated	*Empirical strategy 2 – Matched non-training firms*		
× Roll-out	0.01	-0.02	-0.07
	(0.04)	(0.04)	(0.06)
× Post	-0.11**	-0.07**	-0.16**
	(0.05)	(0.03)	(0.07)
× Phase-out	-0.17**	-0.06	-0.16*
	(0.07)	(0.04)	(0.09)
Observations	6112	5862	6112
R^2	0.02	0.01	0.03
Init. outcome	0.00	0.04	0.00

Notes: Reference group: Treated × Pre. Pre: 1998-2000. Roll-out: 2001. Post: 2002-2004. Phase-out: 2005. Controlling for periods and state fixed effects. Standard errors clustered at the firm level. Init. outcome: Average outcome of treated firms in 1998. For the corresponding event study estimates, see Figure 3.

Table B2.2: DiD Results - Substitution

	# low-educ. trainee hires	# highly educ. commuting trainee hires	Log wages highly educ. trainees	Trainee retention rate	Internal retraining	# low-educ. VT hires	# highly educ. VT hires
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated			*Empirical strategy 1 – All training firms*				
× Roll-out	0.20	0.12	-0.02	-0.11**	-0.01	0.42	-0.18
	(0.71)	(0.17)	(0.04)	(0.05)	(0.05)	(2.56)	(0.28
× Post	-0.39	0.32*	-0.03	-0.09**	-0.09*	3.48	-0.06
	(0.82)	(0.18)	(0.04)	(0.04)	(0.05)	(2.74)	(0.28)
× Phase-out	0.68	0.48**	-0.08*	-0.02	-0.14**	4.87	0.78
	(1.04)	(0.24)	(0.04)	(0.06)	(0.06)	(3.21)	(0.53)
Observations	2296	2295	1758	2260	2227	2296	2296
R^2	0.01	0.01	0.10	0.02	0.02	0.04	0.05
Init. outcome	6.30	0.04	3.00	0.39	0.42	12.46	1.63
Treated			*Empirical strategy 1 – All non-training firms*				
× Roll-out	0.03	-0.04***		0.01	-0.04	-2.75	-0.05
	(0.16)	(0.02)		(0.03)	(0.03)	(2.19)	(0.12)
× Post	-0.23	-0.05***		0.01	-0.08**	-2.29	-0.03
	(0.20)	(0.01)		(0.03)	(0.03)	(2.43)	(0.13)
× Phase-out	-0.25	-0.10***		0.01	-0.04	-3.48	-0.09
	(0.29)	(0.03)		(0.03)	(0.04)	(3.08)	(0.18)
Observations	8382	8373		5798	8178	8382	8382
R^2	0.01	0.01		0.01	0.00	0.01	0.01
Init. outcome	2.34	0.00		0.37	0.39	22.68	1.01
Treated			*Empirical strategy 2 – Matched training firms*				
× Roll-out	0.20	0.26	0.00	-0.12*	-0.05	5.01*	0.05
	(1.04)	(0.33)	(0.04)	(0.07)	(0.06)	(2.99)	(0.40)
× Post	0.02	0.30	0.01	-0.10**	-0.07	4.79	0.05
	(1.10)	(0.24)	(0.05)	(0.04)	(0.07)	(3.37)	(0.34)
× Phase-out	0.01	0.27	-0.05	0.02	-0.12	6.93**	0.71
	(1.50)	(0.32)	(0.05)	(0.07)	(0.11)	(3.46)	(0.62)
Observations	1234	1233	908	1215	1190	1234	1234
R^2	0.02	0.14	0.04	0.04	0.04	0.08	0.03
Init. outcome	5.97	0.03	3.00	0.39	0.43	12.45	1.66
Treated			*Empirical strategy 2 – Matched non-training firms*				
× Roll-out	-0.09	-0.02		0.09*	-0.01	-2.39	0.00
	(0.16)	(0.03)		(0.05)	(0.04)	(2.62)	(0.18)
× Post	-0.25	-0.06**		0.03	-0.07*	-2.01	0.02
	(0.19)	(0.03)		(0.03)	(0.04)	(2.92)	(0.18)
× Phase-out	-0.34	-0.09***		0.06	-0.06	-2.59	0.09
	(0.28)	(0.03)		(0.05)	(0.05)	(4.05)	(0.29)
Observations	5862	5858		3952	5690	5862	5862
R^2	0.02	0.04		0.01	0.01	0.02	0.04
Init. outcome	2.34	0.00		0.37	0.39	22.68	1.01

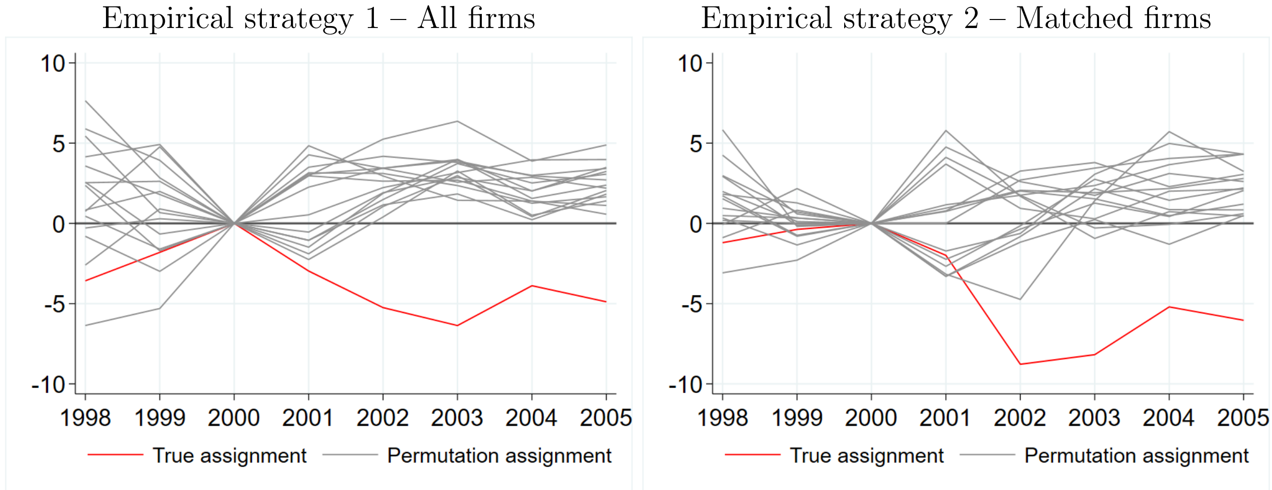
Notes: Full set of results for Table 4. Reference group: Treated × Pre. Pre: 1998-2000. Roll-out: 2001. Post: 2002-2004. Phase-out: 2005. Controlling for periods and state fixed effects. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998.

Table B2.3: 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)
Treated	*Empirical strategy 1 – All training firms*				
× Post	−0.09**	4.93	3.28	3.48	−0.06
	(0.04)	(3.88)	(2.87)	(2.74)	(0.28)
× Phase-out	−0.02	2.42	5.54	4.87	0.78
	(0.06)	(3.93)	(3.51)	(3.21)	(0.53)
Observations	2260	2281	2295	2296	2296
Init. outcome	0.39	23.51	14.04	12.46	1.63
Treated	*Empirical strategy 2 – Matched training firms*				
× Post	−0.10**	5.45	4.43	4.79	0.05
	(0.04)	(7.43)	(3.54)	(3.37)	(0.34)
× Phase-out	0.02	10.22	7.24*	6.93**	0.71
	(0.07)	(7.96)	(3.84)	(3.46)	(0.62)
Observations	1215	1224	1233	1234	1234
Init. outcome	0.39	21.26	14.05	12.45	1.66

Notes: Reference group: Treated × Pre. Pre: 1998-2000. Roll-out: 2001. Post: 2002-2004. Phase-out: 2005. Controlling for periods 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 substituiton effects, see Table 4.

Figure B2.4: Permutation test – Estimation coefficients (Outcome: investments per worker in €1,000)



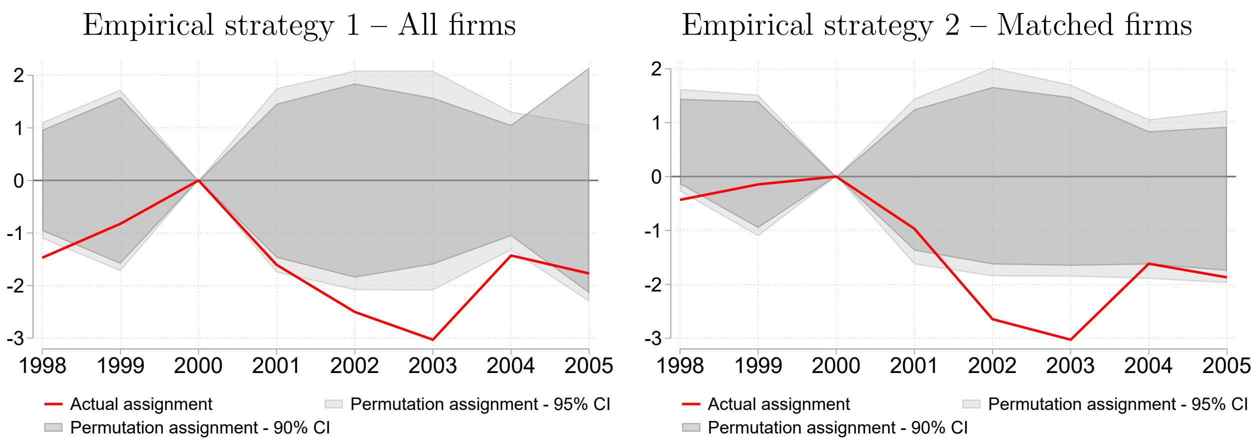
Notes: T-statistics of the event study coefficients of the actual treatment assignment (red line) and all potential permutation assignments within East Germany (gray lines). Outcome: investments divided by total employment in 1998 in €1,000. For the corresponding figures of the t-statistics, see Figure 8.

Table B2.4: DiD Results – Investments

	Investments per worker			Total investments		Ext. margin	Intensive margin		
	Overall	Business serv + Pub. admin.	Manufact.	Absolute	Log	Any inv.	Uppest tercile	Uppest decile	Uppest industry-specific tercile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	*Empirical strategy 1 – All training firms*								
× Roll-out	-1.17	-0.38	-0.24	-306.4	-0.29	0.01	-0.00	-0.02	-0.01
	(1.58)	(2.87)	(2.35)	(508.2)	(0.21)	(0.04)	(0.05)	(0.02)	(0.06)
× Post	-3.37*	-1.86	-3.62**	-1,492.3***	-0.33**	-0.05	-0.07	-0.02	-0.07
	(1.79)	(3.45)	(1.39)	(540.4)	(0.16)	(0.03)	(0.04)	(0.02)	(0.04)
× Phase-out	-3.09	-2.59	0.77	-1,160.1	-0.12	-0.07	-0.02	-0.02	-0.02
	(2.60)	(5.00)	(2.43)	(712.2)	(0.21)	(0.05)	(0.07)	(0.03)	(0.06)
Observations	2344	1040	808	2344	2069	2344	2069	2069	2069
R^2	0.01	0.03	0.03	0.03	0.02	0.03	0.02	0.01	0.02
Init. outcome	17.43	24.22	6.49	6,070.4	7.45	0.89	0.54	0.08	0.35
Treated	*Empirical strategy 1 – All non-training firms*								
× Roll-out	-1.95	-2.93	-0.20	-228.8	0.02	-0.06***	-0.00	-0.00	-0.01
	(1.21)	(3.10)	(1.66)	(155.4)	(0.10)	(0.02)	(0.03)	(0.02)	(0.03)
× Post	-1.53	-1.74	-1.71	-170.5	-0.03	-0.01	-0.02	-0.01	-0.00
	(1.21)	(3.04)	(1.46)	(151.4)	(0.09)	(0.02)	(0.02)	(0.01)	(0.02)
× Phase-out	-0.82	-0.39	-1.78	27.7	-0.12	0.01	-0.00	-0.00	-0.00
	(1.50)	(3.57)	(1.81)	(182.6)	(0.12)	(0.03)	(0.03)	(0.02)	(0.03)
Observations	8744	2376	3792	8744	6881	8744	6881	6881	6881
R^2	0.01	0.03	0.02	0.01	0.01	0.03	0.01	0.01	0.01
Init. outcome	14.87	27.56	8.39	1874.7	5.71	0.83	0.36	0.09	0.33
Treated	*Empirical strategy 2 – Matched training firms*								
× Roll-out	-1.47	-2.80	0.77	-188.2	-0.41*	0.01	-0.03	-0.02	-0.05
	(1.90)	(3.11)	(2.74)	(546.6)	(0.24)	(0.05)	(0.07)	(0.02)	(0.07)
× Post	-6.86***	-10.52**	-3.00	-1,853.8**	-0.50**	-0.03	-0.12***	-0.07**	-0.12**
	(2.29)	(4.04)	(1.88)	(758.12)	(0.21)	(0.05)	(0.05)	(0.03)	(0.05)
× Phase-out	-5.51*	-10.50**	3.05	-927.1	-0.40	-0.02	-0.10	-0.03	-0.11*
	(2.82)	(4.84)	(2.93)	(1,102.0)	(0.29)	(0.07)	(0.08)	(0.04)	(0.06)
Observations	1248	672	336	1248	1068	1248	1068	1068	1068
R^2	0.04	0.06	0.16	0.03	0.05	0.06	0.04	0.03	0.07
Init. outcome	17.68	24.22	6.49	5,736.3	7.41	0.88	0.55	0.09	0.35
Treated	*Empirical strategy 2 – Matched non-training firms*								
× Roll-out	-3.78***	-6.83**	-1.09	-457.6***	-0.10	-0.07***	-0.02	-0.02	-0.02
	(1.35)	(3.20)	(2.10)	(175.7)	(0.13)	(0.03)	(0.03)	(0.02)	(0.04)
× Post	-2.73*	-4.54	-1.64	-200.5	-0.16	-0.03	-0.06*	-0.01	-0.03
	(1.47)	(3.82)	(1.85)	(196.9)	(0.12)	(0.03)	(0.03)	(0.01)	(0.03)
× Phase-out	-2.10	-5.87	-0.36	0.3	-0.16	-0.01	-0.04	-0.00	-0.01
	(1.70)	(4.41)	(1.85)	(228.8)	(0.16)	(0.04)	(0.04)	(0.02)	(0.04)
Observations	6112	1808	2320	6112	4720	6112	4720	4720	4720
R^2	0.02	0.05	0.01	0.01	0.02	0.03	0.02	0.02	0.02
Init. outcome	14.87	27-56	8.39	1,874.7	5.71	0.83	0.36	0.09	0.33

Notes: Reference group: Treated × Pre. Pre: 1998-2000. Roll-out: 2001. Post: 2002-2004. Phase-out: 2005. Controlling for periods 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.

Figure B2.5: 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 - Technological change

			Investment type (0/1)			
	Tech. status	Org. change	Production facilities	ICT	Real estate	Transport *(Placebo)*
	(1)	(2)	(3)	(4)	(5)	(6)
Treated			*Empirical strategy 1 – All training firms*			
× 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)
× 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)
× Phase-out		-0.27	-0.11*	-0.14**	-0.05	-0.07
		(0.17)	(0.06)	(0.06)	(0.05)	(0.06)
Observations	2341	1311	2341	2344	2344	2344
R^2	0.01	0.04	0.05	0.02	0.03	0.04
Init. outcome	3.97	1.35	0.72	0.80	0.59	0.35
Treated			*Empirical strategy 1 – All non-training firms*			
× Roll-out	0.02	0.02	-0.01	-0.07**	0.01	-0.02
	(0.04)	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)
× Post	-0.04	0.03	-0.00	-0.04	0.01	0.01
	(0.04)	(0.06)	(0.02)	(0.03)	(0.02)	(0.02)
× Phase-out		-0.07	0.03	-0.02	0.01	0.03
		(0.07)	(0.03)	(0.03)	(0.03)	(0.03)
Observations	8744	4737	8744	8744	8744	8744
R^2	0.02	0.03	0.02	0.02	0.02	0.01
Init. outcome	3.78	0.91	0.61	0.63	0.35	0.38
Treated			*Empirical strategy 2 – Matched training firms*			
× 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)
× 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)
× Phase-out		-0.20	-0.06	-0.03	-0.01	-0.05
		(0.24)	(0.09)	(0.07)	(0.08)	(0.10)
Observations	1245	702	1248	1248	1248	1248
R^2	0.03	0.04	0.07	0.06	0.11	0.09
Init. outcome	3.98	1.41	0.71	0.79	0.58	0.33
Treated			*Empirical strategy 2 – Matched non-training firms*			
× Roll-out	0.01	0.04	-0.01	-0.06	0.00	-0.07**
	((0.05)	(0.08)	(0.04)	(0.04)	(0.03)	(0.03)
× Post	-0.04	0.04	-0.02	-0.05	-0.03	-0.03
	(0.06)	(0.09)	(0.03)	(0.04)	(0.03)	(0.03)
× Phase-out		0.03	0.00	-0.01	-0.03	0.00
		(0.09)	(0.04)	(0.05)	(0.04)	(0.04)
Observations	6112	3308	6112	6112	6112	6112
R^2	0.02	0.04	0.03	0.03	0.03	0.01
Init. outcome	3.78	0.91	0.61	0.63	0.35	0.38

Notes: Full set of results for Table 6. Reference group: Treated × Pre. Pre: 1998-2000. Roll-out: 2001. Post: 2002-2004. Phase-out: 2005. Controlling for periods 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 B2.6: DiD Results - Other outcomes

	(1) Sales per worker	(2) Log employment	(3) Log wages	Firm exit
Treated	*Empirical strategy 1 – All training firms*			
× Roll-out	-37.78	-0.05*	0.01	-0.013
	(23.64)	(0.03)	(0.01)	(0.011)
× Post	-2.79	-0.11**	-0.00	-0.002
	(26.06)	(0.05)	(0.01)	(0.006)
× Phase-out	12.35	-0.06	0.01	0.004
	(37.69)	(0.08)	(0.02)	(0.005)
Observations	1260	2344	2344	8894
R^2	0.05	0.04	0.12	0.00
Init. outcome	234.83	5.21	4.17	
Treated	*Empirical strategy 1 – All non-training firms*			
× Roll-out	-13.38**	-0.01	-0.00	0.009
	(6.03)	(0.02)	(0.00)	(0.006)
× Post	-15.32**	-0.03	0.00	0.010**
	(7.45)	(0.02)	(0.01)	(0.004)
× Phase-out	-7.95	-0.03	-0.00	0.006
	(9.06)	(0.03)	(0.01)	(0.004)
Observations	5972	8744	8744	36396
R^2	0.02	0.02	0.06	0.00
Init. outcome	147.45	4.16	3.99	0.022
Treated	*Empirical strategy 2 – Matched training firms*			
× Roll-out	-40.70	-0.04	0.01	-0.006
	(27.98)	(0.03)	(0.01)	(0.007)
× Post	-32.96	-0.06	-0.01	-0.006
	(35.34)	(0.06)	(0.01)	(0.006)
× Phase-out	2.02	0.05	0.02	0.002
	(53.59)	(0.12)	(0.02)	(0.006)
Observations	558	1248	1248	7979
R^2	0.08	0.06	0.09	0.01
Init. outcome	245.37	5.18	4.17	0.00
Treated	*Empirical strategy 2 – Matched non-training firms*			
× Roll-out	-15.77**	-0.00	0.00	0.010**
	(6.77)	(0.02)	(0.00)	(0.005)
× Post	-17.94**	0.00	-0.00	0.003
	(8.24)	(0.03)	(0.01)	(0.006)
× Phase-out	-15.75	0.01	-0.01	0.008*
	11.69	(0.05)	(0.01)	(0.004)
Observations	4046	6112	6112	34419
R^2	0.02	0.02	0.06	0.01
Init. outcome	147.45	4.16	3.99	0.00

Notes: Reference group: Treated × Pre. Pre: 1998-2000. Roll-out: 2001. Post: 2002-2004. Phase-out: 2005. Controlling for periods 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		✓		✓		✓		✓

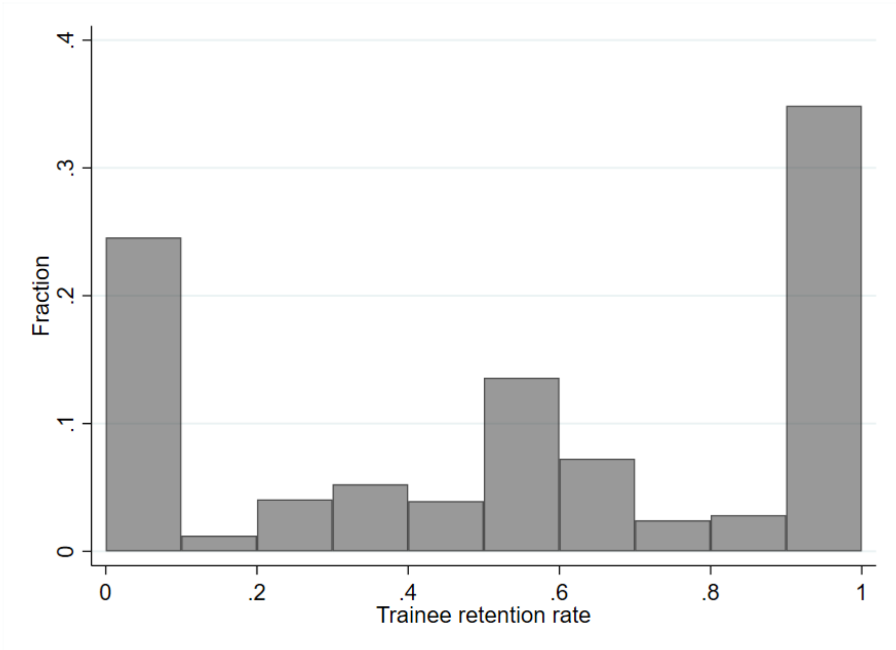
Notes: Outcome: Column 1–4: investments per worker in €1,000. Columns 5–8: Technical status of machinery on a scale from 1 "completely-out-of-date" to 5 "state-of-the-art". To avoid confusion by the education reform, including firms in control states only. Controls include year fixed effects, industry, 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	0.04	-0.13	-0.13	-0.12	-0.12	-0.08
	(0.05)	(0.14)	(0.13)	(0.13)	(0.13)	(0.11)
Industry		✓	✓	✓	✓	✓
Exposure			✓	✓	✓	✓
Firm size				✓	✓	✓
Definition of new skills based on						
... Years	2000-2001	2000-2001	2000-2001	2000-2001	1998-2001	2000-2001
... Worker type	All	All	All	All	All	Trainees
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 9.

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 9.

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 9.

C Instrumental variable regression

I examine the treatment effect along the intensive margin of the negative trainee supply shock using a complementary identification strategy. This sheds light on the between-firm effect and strengthens the argument that the investment declines are indeed caused by the negative trainee supply shock. In addition, it provides an estimate of the investment decline associated with each absent highly educated trainee.

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.

To disentangle trainee supply from demand, and to estimate the effect of one additional trainee supplied, I next estimate the following two-stage-least-squares (2SLS) model:

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

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

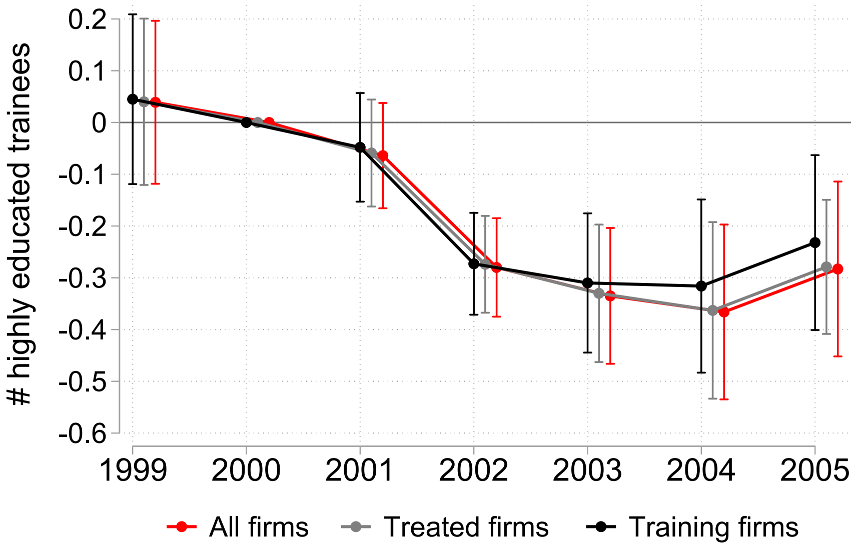
with equation (C2) capturing the first-stage equation and equation (C1) the second stage. Inv denotes investments, N^{Trainee} employment of highly educated trainees, j firm, b federal states, and t calendar year. Treated is a binary variable with Treated $= 1$ if the firm is located in a state undergoing 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, $N^{\text{Trainee}}_{1998}$, 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. As expected, 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.

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 in-

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.

vestment 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.

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

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	572.0***	228.0*	551.2**	87.1	600.1**
	(51.2)	(216.8)	(121.6)	(217.9)	(55.4)	(235.8)
Observations	9702	9702	3241	3241	2051	2051
p-value KP		0.025		0.019		0.061
F-Stat		22.24		23.34		14.53

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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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, again, suggests spill-over effects among firms with no or few highly educated trainees.

D Economic framework

In this Appendix, I provide a more formal exposition of the economic reasoning why labor market entrants act as complements to new technologies. I follow Ahituv & Zeira (2011) in the way technological progress is introduced but model the firm optimization problem.

Baseline Setting. Suppose that firms operate and employees work in overlapping generations for an infinite amount of periods T.[33] In each period $t \in T$, each firm j produces one final good using labor L and production technologies τ with fixed marginal productivities a_τ under the following production function:

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

Technologies require vintage-specific skills, i.e. only workers trained for a specific technology, L_τ, can handle this technology. Except for workers ability to handle technologies, workers are homogenous. Training takes one period uniformly across technologies and workers. At the beginning of each period, a new 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 output per worker 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. To adopt, firms (re-)train a fraction Ψ_τ of workers of each cohort. Since workers within a cohort are homogeneous, firms always either retrain all or no worker of one cohort, $\Psi_\tau = \{0; 1\}$. Costs of adopting the new technology consist of one-time capital costs K_τ, and costs for worker training, which are equal to foregone production output remunerated at the workers' respective wage levels. 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.[34]

For simplicity, assume that workers do not switch firms, and that firms maximize profits with respect to the current period, $t = 1$, and next period, $t = 2$ only. This assumption is relaxed to an infinite time horizon below. Additional profits from adopting the new technology τ, $\Delta\pi_{j\tau}$, are equal to the net surplus in output minus wages in the next period, $\Delta Y_{j\tau} - \Delta W_{j\tau}$, minus training costs $\overline{Y_j}$, and capital costs K_τ:

$$\Delta\pi_{j\tau} = \Delta Y_{j\tau} - \Delta W_{j\tau} - \overline{Y_j} - K_\tau \tag{D2}$$

[33] Allowing for worker retirement gives an additional reason to train young workers over incumbent workers – a channel I want to abstract from.

[34] 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).

Net surplus in output minus wages are equal to the sum of productivity increases minus wage increases across all initial types L_{τ_0} trained in the new technology (D3). Training wages are equal to the sum of foregone outputs of all retrained workers (D4):

$$\Delta Y_{j\tau} - \Delta W_{j\tau} = (1-\theta) \sum_{\tau_0=0}^{\mathcal{T}-1} \Psi_{\tau_0} L_{jt\tau_0} (a_\tau - a_{\tau_0}) \tag{D3}$$

$$\overline{Y_j} = \sum_{\tau_0=0}^{\mathcal{T}-1} a_{\tau_0} \Psi_{\tau_0} L_{jt\tau_0} \tag{D4}$$

Firms maximize $\Delta\pi_{j\tau}$ by choosing whether or not to (re-)train each initial worker type. The net output surplus increases the lower the initial productivity of the worker, while training costs decrease the lower the initial productivity. Firms hence prefer to train their least productive workers, i.e. their least trained workers. In particular, combining equations (D2)–(D4), it follows that firms train a worker type L as long as the following conditions between initial productivity a and the productivity of the new technology, a_τ, holds:

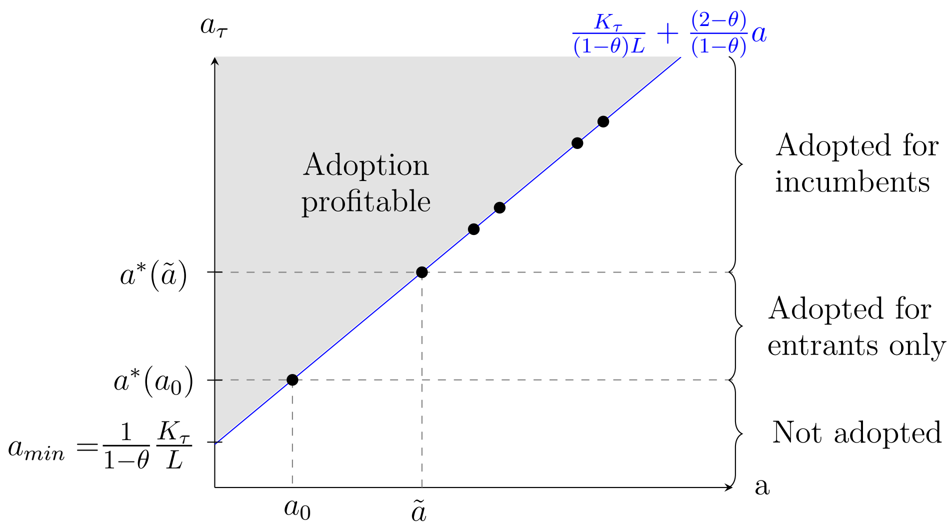
$$a_\tau \geq \frac{K_\tau}{(1-\theta)L} + \frac{(2-\theta)}{(1-\theta)} a \tag{D5}$$

Figure C2 visualizes the adoption indifference curve. Technologies below the productivity threshold a_{min} are never adopted because their capital costs K are too high. Technologies above a_{min} but below the threshold $a^*(a_0)$ would be adopted if workers with an initial productivity below a_0 were available. Technologies above $a^*(a_0)$ but below $a^*(\tilde{a})$ are adopted by training labor market entrants with initial productivity a_0 only, while technologies above the threshold $a^*(\tilde{a})$ are adopted for workers with higher initial productivity levels as well. Due to the training costs, the gradient of the indifference curve is larger than one. Hence, in order to make retraining of a twice as productive worker profitable, the new technology has to be more than twice as productive.

Negative supply shock of trainees. Assume there is a missing entry cohort in $t = 2$. In this period, firms will invest in the new technology if and only if the productivity gain is large enough to make it profitable to retrain incumbent workers. For productivity levels of the new technology $a^*(a_0) \leq a_\tau < a^*(\tilde{a})$, this implies a reduction in firms' technology adoption compared to the case without a missing entry cohort. Note that missing entrants can only be substituted when workers from previous cohorts are still untrained because all past technologies were not productive enough to pay off costs of training ($a_\tau < a^*(a_0)$).

Extension A – Infinite time horizon and worker retention. I now account for the fact that workers might leave firms and firms maximize expected profits of all future periods. For simplicity, assume there is no discounting. The expected total surplus of adopting a new

Figure C2: Indifference curve



technology τ is now given by the sum of all expected future output increases net wage increases, minus one-time training and capital costs:

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

Workers can leave 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. 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.[35] For the purpose of this study, the underlying reasons are irrelevant, and I assume p_j to be exogenously given.

For each worker group, the net outplus surplus from adoption extincts as soon as this worker group is retrained in a new technology. I annotate the probability to be retrained $\pi(p_j)$, with π increasing in p_j. Hence, the expected total net output increase is equal to the net output surplus of each worker in each (future) period multiplied by the probability of still being at the firm in this period and still not being (re-)trained:

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

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

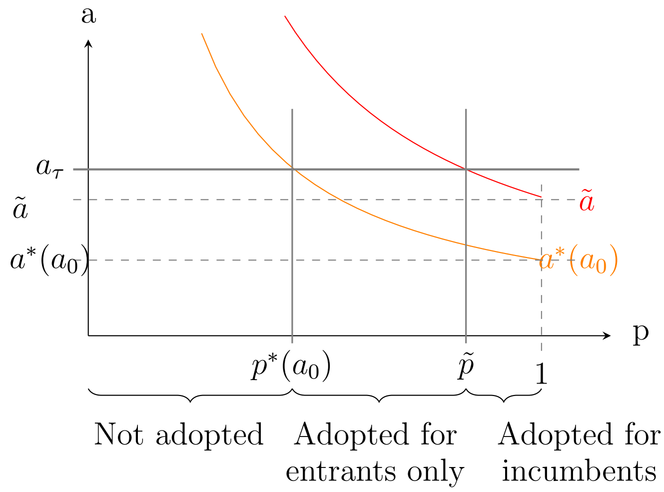
$$a_\tau \geq \frac{K_\tau}{L}\frac{1}{1-\theta}\frac{1}{\sum_t p^t(1-\pi(p_j))^t} + \left(1 + \frac{1}{1-\theta}\frac{1}{\sum_t p^t(1-\pi(p_j))^t}\right)a \tag{D8}$$

Figure C3 visualizes the indifference curve along p for entrants with the productivity a_0 and

[35] See the excellent survey by Wolter & Ryan (2011).

incumbents with the productivity $\tilde{a}$. Technologies below the productivity threshold a_{min} are never adopted because their capital costs K are too high. A technology of productivity a_θ is not adopted for retention rates below $p^*(a_0)$. For retention rates above $p^*(a_0)$ but below $\tilde{p}$, the technology is adopted for entrants only. For retention rates above $\tilde{p}$, the technology is adopted for incumbents as well. Given a retention rat p, training is more more profitable the lower the initial productivity of the worker.

Figure C3: Indifference curve 2



Extension A + Negative supply shock of trainees. Let us turn to the case when no new, untrained cohort L_0 is 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 a new, untrained cohort, because they would not have adopted in the counterfactual scenario either. For firms with a retention rate above $p^*(a_0)$ but below $\tilde{p}$, technology adoption is lower than in the counterfactual scenario. For firms with a retention rate above $\tilde{p}$, 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.

Extension B – Technologies not requiring new skills. Assume that some of the new technologies arriving each period do not require any retraining for exogenous reasons. Costs of adopting these technologies consist of capital costs only.

In the two-period setting, firms train a worker type L as long as the following condition holds:

$$a_\tau \geq \frac{K_\tau}{(1-\theta)L} + a \tag{D9}$$

In the setting with an infinite time horizon, firms train a worker type L as long as the following condition holds:

$$a_\tau \geq \frac{K_\tau}{L} \frac{1}{1-\theta} \frac{1}{\sum_t p^t (1-\pi(p_j))^t} + a \tag{D10}$$

In both cases, training incumbent workers is still less profitable than training entrants because productivity gains of training are higher. However, since costs of adoption do not increase with initial productive, the range of technologies which are adopted for entrants only decreases. In the two-period setting, this corresponds to a flatter indifference curve in Figure C2. In the infinite time horizon setting, the two indifference curves move closer to each other. In the absence of an entry cohort, the range of technologies that is not adopted decreases.