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

Cäcilia Lipowski*

[Link to latest version]

September 10, 2024

Abstract

Firms in developed countries are increasingly facing a declining supply of young workers. This paper studies the relevance of young workers, particularly vocational trainees, for firms' technology investments. Leveraging exogenous variation in the supply of trainees caused by an education reform in Germany in 2001, I show that a reduction in the supply of trainees decreases firm technology investments, implying complementarity between young workers and new technologies. This is explained by young labor market entrants having lower opportunity costs and higher returns to learning new tech skills than incumbent workers.

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

JEL: D22, D24, J21, J24, O33

---

*ZEW Mannheim, caecilia.lipowski@zew.de

I am greatly indebted to Anna Salomons and Ulrich Zierahn-Weilage for their advice and support. For helpful comments and conversations, I thank Daron Acemoglu, Melanie Arntz, David Autor, Eduard Brüll, Christian Dustmann, Katja Görlitz, Maarten Goos, Simon Jäger, Morten Olsen, Harald Pfeifer, Pascual Restrepo, Johannes Schmieder, Anna Waldman-Brown, and Nicolas Ziebarth. I thank conference and seminar participants at the EEA (Barcelona), EALE (Prague), TPRI (Boston), Skills-for-the-Future Conference (LISER), ces-ifo Summer Institute on Re-skilling and Skill Shortages (Venice), Leading House Conference on the Economics of Educational and Vocational Training (Zurich), and at internal seminars at University of Bonn, Frankfurt School of Finance and Management, ifo Munich, IWH Halle, IZA, LMU Munich, University of Mannheim, Rockwool Foundation Berlin, 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

Young labor market entrants likely matter for firms' technology investments: compared to incumbent workers, young labor market entrants can acquire new tech skills at lower opportunity costs, they have higher expected returns to human capital investments, they may inherently possess more digital skills, and they can be hired in occupations demanded for technological transitions.[1] Reduced supply of young labor market entrants may therefore reduce firms' technology investments. In contrast, classic economic theory predicts that supply reductions of young labor market entrants increase firms' incentives to adopt labor-saving technologies. Understanding how the supply of young labor market entrants affects firms' technology investments is crucial in times of accelerated technological change and population aging. The identification of the causal effect is challenging because changes in the supply of young labor market entrants usually go hand in hand with changes in the supply of other types of labor as well as changes in labor demand; and are often confounded by unobserved factors at the region, industry, or firm level. Empirical evidence on the causal effect is therefore lacking.

In this paper, I overcome this identification issue and provide empirical evidence on the causal effect of a supply reduction of young labor market entrants on firms' technology investments. I exploit a natural experiment created by an education reform. In 2001, two out of six East German federal states, henceforth "treated states", permanently increased the length of schooling required for the university entrance qualification by one year,[2] causing a missing school graduation cohort from the upper school track. There was no comparable reduction in the other four East German states, henceforth "control states". Since labor market entry in Germany often occurs via a three-year-long vocational training, the missing school graduation cohort translated into one missing trainee entry cohort and significantly reduced the supply of trainees in the subsequent three years in treated states. The missing trainees from the upper school track can be thought of as currently unskilled but future middle-skilled professionals, who make up a large share of a firm's "transformative" workforce: 12% of a training firm's hires and 13% of a training firm's young workers below the age of 30. They often work in white-collar occupations such as media, retail, or financial service occupations, which commonly require bachelor's or associate degrees in other countries like the US.

I compare investments and technology adoption of firms in treated East German states undergoing the temporary trainee shortage to investments and technology adoption of firms in control East German states in a difference-in-differences event study design. To ensure that no concomitant industry-specific shocks drive the results, I ensure the comparability of treated and control firms by matching treated firms to comparable control firms operating in the same

[1] See MacDonald & Weisbach (2004); Aubert et al. (2006); Ahituv & Zeira (2011); Cavounidis & Lang (2020); Deming & Noray (2020); Adão et al. (2024).

[2] Among others, Büttner & Thomsen (2015); Morin (2015); Muehlemann et al. (2022); Marcus & Zambre (2019) and Dorner & Görlitz (2020) exploit this and the opposite reform to study the effect on school grades, university enrollment, trainee employment and trainee wages. So far, no study has looked at effects on firms.

sector. I focus on training firms, defined as firms that employed trainees from the upper school track prior to the reform. Non-training firms should not be directly impacted by the shock and serve in a falsification test.

This setting provides a unique opportunity to study the causal effect of supply reductions of young labor market entrants on firm technology adoption because it provides plausibly exogenous variation in the supply of trainees across firms and time. The low geographic mobility of trainees greatly enhances the sharpness of the negative trainee supply shock with respect to state boundaries, aiding identification.[3] In addition, the supply shock is devoid of a potentially concomitant demand shock.[4] Also, the shock affects a large part of the labor market: two-thirds of the German workforce hold a vocational training degree.[5] At the same time, the implications of the natural experiment can likely be extended to other settings, though with caution, since the proposed mechanism is expected to hold beyond the specific context.

I use a large and representative firm panel survey linked with social security records that allows me to directly observe trainee employment, investments in tangible assets plus ICT (information and communication technologies), and the technical status of a firm's machinery. While the data lacks information about the specific investment made or technology adopted, it comes with the advantage of encompassing a broad spectrum of investments and technologies rather than concentrating solely on one such as robots or computers.

I provide three key empirical findings. First, the education reform produces trainee shortages. The reform has a substantial negative effect on firms' employment of trainees from the reformed school track, i.e. trainees with 12 or 13 years of schooling and a university entrance qualification, henceforth "highly educated trainees". Highly educated trainees make up 16% of all trainees (Federal Statistical Office, Genesis-Online, 2022a) while the majority only have 9 or 10 years of schooling, henceforth "low-educated trainees." Training wages do not increase, likely due to wage rigidities and the shock being temporary. Since trainees tend to be very geographically immobile, there is no increased commuting of trainees across states. Given their differences in education levels and usual occupation choices, firms do not compensate for missing highly educated trainees by hiring more low-educated trainees. Workers who have completed their training program are not hired to compensate for the missing trainees either and there is not expansion of internal training for incumbent workers. This suggests that already trained workers are no good substitutes for young labor market entrants.

The second key finding is that trainee shortages cause reductions in firm investments: investments decrease sharply in training firms in treated states compared to control states in the face of the trainee shortage. This finding addresses the central question this paper raises:

[3]Only 2.2% of trainees move across federal states for their vocational training (Socio-Economic Panel (SOEP), own calculations). Likewise, only 5% of trainees commute between federal states (LIAB, own calculations).

[4]The number of consumers is not affected by the shock, only the composition: some consumers are now trainees instead of students. Since trainees earn very low wages, their consumption behavior is little affected.

[5]Sample of Integrated Labor Market Biographies (SIAB), own calculations.

Trainees and investments are complements rather than substitutes, and their scarcity does not induce firms to invest more in order to compensate for their absence but rather impedes investments. In line with the notion of complementarity, firm investments of treated firms catch-up with those of control firms once the shock is over. At the same time, the technologies (not) adopted at the time are unlikely to be complements to labor in general because labor in general was not scarce, yet investments decrease. The effect is large: investments temporarily drop by approximately one-fifth of a standard deviation in affected years. The sizable decrease in investments is driven by firms refraining from large investments, in line with the fact that investments are lumpy. Based on a decomposition, I show that the decrease in investments is not mainly driven by a decrease in firm employment size. I confirm the link between the investment decline and the absence of trainees in two ways. First, comparing non-training firms across treated and control states, I do not find a comparable reduction in investments. Second, employing an auxiliary identification strategy, i.e. a Bartik-type instrument exploiting pre-reform exposure to the shock, I show that firms which are more affected by the negative trainee supply shock decrease investments to a greater extent than less affected firms.

The third key finding is that the induced investment decline is linked to reduced adoption of new technologies: the technical status of machinery depreciates in treated training firms compared to control training firms once foregone investments accumulate. 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).

The complementarity between young labor market entrants and new technologies can be rationalized through the lens of a model of endogenous technological change in which new technology vintages create new tasks that demand new vintage-specific skills. Firms assign these new tasks to workers who have the lowest cost of acquiring new skills. Compared to incumbent workers, opportunity costs of training new skills in terms of foregone output are low for young, initially unproductive, labor market entrants, and concomitant productivity gains of training are large. Firms thus choose to complement their technology adoption with young labor market entrants.

Take for example the advanced office technologies adopted around the time of the education reform. Industrial clerks use them to manage orders, inventory, production schedules, and customer interactions.[6] These new technologies required new skills, as can be seen from the fact that the vocational training curriculum of the industrial clerk was changed in 2002. Supporting the idea that mainly young workers acquire the new skills, more young workers worked with these new technologies: in 2006, 22% of office clerks below the age of 30 reported to use new technologies, but only 17% of workers aged 30 and above.[7]

I provide three pieces of empirical evidence in support of the mechanism via technology-

[6] These technologies include software such as Microsoft Outlook which became essential for task coordination, Enterprise Resource Planning (ERP) systems like SAP and Oracle, Customer Relationship Management (CRM) systems like Salesforce, and Supply Chain Management (SCM) software.

[7] See Section 7.2.

vintage specific skills: First, a heterogeneity analysis in the setting of the educational reform reveals that the investment drop is large among firms where incumbent workers have outdated skills, and small in firms where incumbent workers possess up-to-date skills. Second, in a firm survey, training firms largely agree that vocational training ensures the supply of new skills and helps adapt to technological change. Third, based on an employee survey, young workers work significantly more often with new technologies than older workers. This pattern is found for young workers of any education, indicating that the overall finding—young labor market entrants and technology adoption being complementary—can likely be extended to settings beyond the trainee supply shock studied in this paper. Alternative explanations for the complementarity between young labor market entrants and technology adoption, for example via trainees generally possessing more up-to-date skills than incumbents, are unlikely to explain the bulk of the effect.

This paper contributes to four literatures. The most closely related strand of literature studies how technology invention and adoption respond endogenously to the relative abundance of production factors (e.g. Zeira, 1998; Acemoglu, 1998, 2002). Empirical papers, mainly exploiting migration shocks, find support for this theory. For example, a decrease in the supply of *low*-skilled labor increases labor-saving patenting and fosters the adoption of labor-saving production technologies (Lewis, 2011; Hornbeck & Naidu, 2014; Clemens et al., 2018; Dechezleprêtre et al., 2019; Danzer et al., 2020; Andersson et al., 2022; San, 2023, also vice versa for an increase). In turn, an increased supply of *high*-skilled labor intensifies the adoption of skill-complementing technologies (Beaudry et al., 2010; Carneiro et al., 2022). This paper shows that the supply of young labor market entrants is a key ingredient for firm technology adoption, hereby contributing to the literature on endogenous technological change in two dimensions. First, it studies the role of young labor market entrants—a decisive worker group that has thus far been ignored in the literature,[8] and highlights a key characteristic of young labor market entrants; the comparatively low opportunity costs of training them. Second, it highlights the role of capital adjustment costs of worker training for investment decisions—a factor deliberately ignored in models about endogenous technological change. If technological change is endogenous to factors entering capital adjustment costs, this can produce substantially different results than standard models.

Second, I provide strong support for the relevance of vintage-specific skills, hereby contributing to a literature demonstrating how new technologies require new skills and create new tasks (e.g. Chari & Hopenhayn, 1991; Autor et al., 2003; Acemoglu & Restrepo, 2018; Autor et al., 2024; Lipowski et al., 2024). Such new skills have been linked to decreasing returns to worker experience, and reduced employment of older workers (e.g. Aubert et al., 2006; Ahituv & Zeira, 2011; Deming & Noray, 2020). They have also been put forward as the reason why

[8]An exception are macro papers on endogenous technological change and demographic change: countries with lower population growth or shortages of middle-aged workers are found to 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).

adaptation to technological change takes place through the entry of young workers, rather than by upskilling incumbent workers (MacDonald & Weisbach, 2004; Cavounidis & Lang, 2020; Adão et al., 2024). This paper manifests that technology vintage-specific skills hinder technology adoption when young labor market entrants are scarce.

Third, this paper relates to the literature on firm training, which has highlighted various aspects of why firms may or may not provide training (Becker, 1962, 1964; Acemoglu & Pischke, 1998, 1999a,b; Moen & Rosén, 2004; Dustmann & Schönberg, 2009, 2012; Caicedo et al., 2022). This paper demonstrates that retraining incumbents is costly, incentivizing firms to train young labor market entrants to keep pace with technological advances.

Finally, I contribute to nascent literature on the consequences of labor shortages on firm outcomes. While existing studies establish a negative effect on firm capital, sales, and productivity (D'Acunto et al., 2020; Le Barbanchon et al., 2023; Sauvagnat & Schivardi, 2024), I provide detailed evidence on one mechanism through which reduced labor supply affects firm outcomes, namely technology adoption.

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

2 The German vocational training system and the education reform

Below, I describe the functioning of the German vocational training system and detail the education reform used for identification.

2.1 The German vocational training system

Vocational training is a key component of both the German education system and labor market, with approximately 60% of the working population having undergone such training (Sample of Integrated Labour Market Biographies, own calculations). In the context of this paper, vocational trainees can be regarded as yet unskilled individuals with their single purpose being to acquire skills and become middle-skilled professionals. Vocational training often prepares individuals for occupations that typically require bachelor's or associate's degrees in other countries, such as the US.

Adolescents usually start vocational training after graduating from one of the following three high-school tracks: the basic track (*Hauptschule*, 9 years of schooling) which qualifies students for vocational training in blue-collar occupations; the intermediate track (*Realschule*, 10 years) which prepares students 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 choose to undergo vocational training,[9] such that in 2000, 16% of trainees had a university entrance qualification (*Abitur;* Federal Institute for Vocational Education and Training, 2009). Trainees rarely move or commute to their workplace: only 2% of vocational trainees move across states for their vocational training (SOEP, own calculations). Based on the data used in the subsequent analyses, the share of trainees commuting across states is similarly low at approximately 5%. After completing the vocational training, which typically lasts three years, a high share of trainees remain at their training company. Trainees from the upper school track often work in media, financial services, or retail occupations, but are also found in manufacturing and technical occupations.

Vocational training in Germany 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). This paper exclusively focuses on the on-the-job training part. Trainees are hired by their training company, receive a work contract for the duration of their vocational training, and are paid a training wage, even though training wages are usually subject to collective bargaining agreements and are low.[10] Regarding the central aspects of this paper, vocational training is comparable to on-the-job training in other countries with two notable exceptions: First, trainees receive state-provided vocational schooling in addition to training at the firm. Second, nationally binding training curricula ensure that the training content is both current and not firm-specific.

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. Saxony-Anhalt and Mecklenburg-Western Pomerania transitioned from 12 to 13 years with the graduation cohort of 2001. This switch constitutes the source of the shock that I exploit in this paper. In what follows, I therefore assign Saxony-Anhalt and Mecklenburg-Western Pomerania as treated states and the other

[9]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.

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

four East German states as control states. The education reform was decided in May 1996 in Mecklenburg-Western Pomerania and in January 1998 in Saxony-Anhalt.[11] 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 reform resulted in a missing upper-track school graduation cohort in spring 2001. Figure 1, Panel A depicts the 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 other East German states, from now on referred to as control states.

How does the missing school graduation cohort translate into the labor market? Usually, two-thirds of the missing upper-track school graduates eventually opt for university studies, while one-third eventually start vocational training. The missing school graduates of spring 2001 are hence expected to result in a missing entry cohort of highly educated trainees in fall 2001, and to reduce the stock of highly educated trainees for three consecutive years given that vocational training typically lasts three years. At that time, males in Germany had to do military service of 10 months when reaching the age of 18, partly postponing the missing entry and the reduction in the supply of trainees by one year.

Official statistics confirm this decline in the supply of trainees: Figure 1, Panel B shows that training contracts with school graduates from the upper-track evolved in parallel in treated and control states between 1998 and 2000 but sharply fell in 2001 and 2002. While the decline is meaningful in magnitude, new training contracts with upper-track school graduates do not decrease by 100% but rather by approximately 28%,[12] suggesting that school graduates from other states, or other school graduation cohorts partly compensate for the shock.

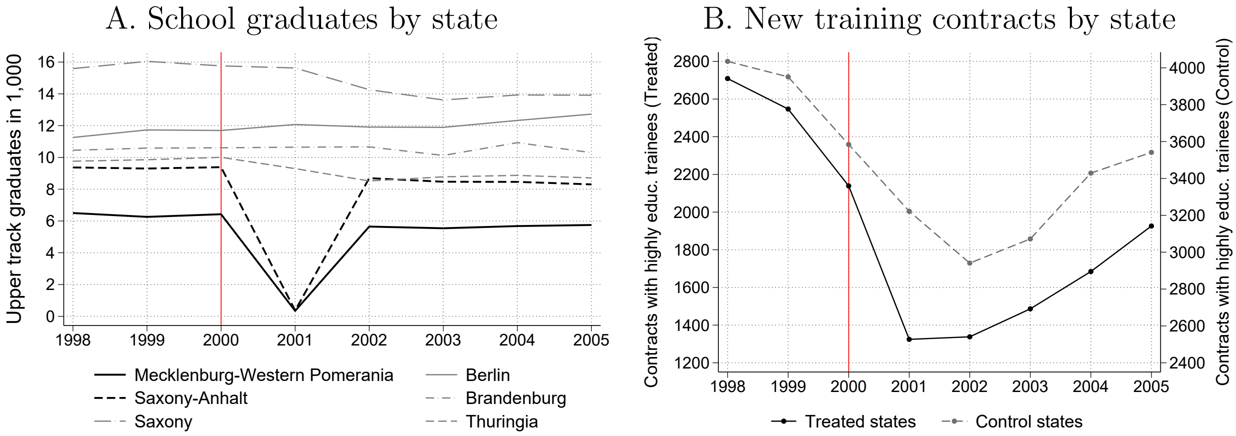
This is the labor supply shock in upper-track vocational trainees I exploit for identification. How the education reform reduced trainee employment is studied in detail in Dorner & Görlitz (2020). I instead use this shock to trainee supply as the first stage, to study subsequent effects on investments. I focus on upper-track school graduates who subsequently start vocational training instead of university students/graduates because vocational trainees postpone their labor market entry less and move or commute less across federal states, thus endorsing the credibility of the identification strategy. Note that the labor supply shock is unlikely to be confounded by a labor demand shock: the overall number of consumers remains unchanged, only the composition adjusts. Since trainees earn low wages, consumption patterns of trainees

[11]For more information on the education reforms, see Kühn et al. (2013) and 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. To avoid potential confounding effects from these changes, this study ends in 2006.

[12]This number can be calculated in two different ways. First, in 2001 and 2002, 717 training contracts less are concluded on average in each treated state compared to 2000, while approximately 2,500 school graduates are missing in each treated state, corresponding to a drop by $717/2,500 = 28\%$ of the school graduates. Second, the training contracts in treated states in 2001 dropped by 38% compared to 2000, and by 10% in control states, suggesting a reform-induced decline by 28% of previous training contracts.

versus students is unlikely to have caused relevant demand changes.

Figure 1: The missing school graduation cohort



Notes: Panel A: Total number of upper-track school graduates per federal state. Source: Federal Ministry of Education & Research (2022). *Panel B:* Average number of new training contracts within the dual system with graduates from the upper school track across treated states (Mecklenburg-Western Pomerania and Saxony-Anhalt) and control states (Berlin, Brandenburg, Saxony, Thuringia). Source: Federal Statistical Office, Genesis-Online (2022a).

Perhaps improving the credibility of the research design, both treated states are economically fairly different: Mecklenburg-Western Pomerania, located in the northeast of Germany along the Baltic Sea, is a predominantly rural and sparsely populated federal state with approximately 1.6 million inhabitants as of 2020. Its economy is defined by small and medium-sized enterprises engaged in agriculture, maritime industries, mechanical engineering, and tourism. Saxony-Anhalt, situated in central Germany with a population of around 2.2 million in 2020, features a comparatively more urban environment. It is characterized by the chemical industry, mechanical engineering, and automotive supply. Both states, as well as control states, are characterized by excess trainee supply and high unemployment rates during this period, namely 17.8% in Mecklenburg-Western Pomerania and 20.2% in Saxony-Anhalt in 2000, see Figure C1, Panel B.

The education reform was a claim of the Social Democratic Party, which entered the government in both treated states in 1994. I rule out that the governance of the Social Democrats, or related policy or socio-economic changes, confound the effect of the education reform by comparing several state metrics between treated and control states before and after the reform, as well as between state-periods governed by the Social Democrats and those not governed by the Social Democrats, in Appendix B. In Appendix B, I also discuss why concomitant investment tax programs are unlikely to have confounded the trainee supply shock.

3 Firm panel data

The analysis is based on the Linked-Employer-Employee-Data of the Institute for Employment Research (IAB), the LIAB-QM, which combines the annual representative IAB Establishment Panel survey with administrative employment information of all employees at surveyed firms.[13] 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 provide a reliable distinction between trainees and workers who have completed their training program, 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 qualification and those with a lower schooling degree, respectively. This is important since the education reform directly affects highly educated trainees only.[15]

I restrict the data in the following ways. First, I focus on the period 1997–2006. Second, I limit the data to firms in East Germany including Berlin, since the firms in East Germany are likely not comparable to firms in West Germany, especially for this period relatively shortly after reunification. Third, I exclude firms in the health/education/social service sectors because vocational training in many related occupations is entirely school-based. Fourth, I drop very small firms, defined as those that never reach ten employees, because they tend to exhibit volatile training investment behavior. Fifth, I only keep observations with non-missing investment values. Sixth, I only keep firms that have existed in 1997. There is panel attrition in firms participating in the survey, see Figure A1: 32% of firms observed in 1997 are still present in 2006. Estimates for later years must therefore be interpreted with caution.

The final sample comprises 2,303 distinct firms, of which 775 are treated (397 in Saxony-Anhalt and 346 in Mecklenburg-Western Pomerania) and 1,528 are untreated. Table 1 shows summary statistics of the final dataset. In sum, all firms cover more than 280,000 workers per year, amounting to approximately 3.9% of the East German workforce in a year.[16] I observe 15,681 trainees on average across years, of which 2,541 (16%) are highly educated. In 78% of

[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.

[15]I use the harmonized version of the schooling variable based on the imputation procedure by Thomsen et al. (2018) and Fitzenberger et al. (2006).

[16]The average yearly working population in East Germany from 1997 to 2006 was 7.43 million according to Statistisches Landesamt (2023).

the firm-by-year observations, no highly educated trainee is employed, and 61% of firms never employ a highly educated trainee over the entire time window 1997–2006.

Table 1: Summary statistics

	Mean	SD	Min	Max	Yearly sum
# workers	171.02	448.44	1	12133	286,793
# trainees	9.35	51.52	0	3181	15,681
# highly educated trainees	1.51	8.78	0	461	2,541
No highly educated trainee	.78	.42	0	1	1,301
No highly educated trainee ever (1997–2006)	.61	.49	0	1	1,021

Notes: Summary statistics in the full data sample (training and non-training firms, 1997–2006). SD: standard deviation. Yearly sum: Sum of workers across all observed firms.

Table 2: Pre-reform averages among training versus non-training firms

	Training firms	Non-training firms	Δ
# workers	433.96	105.15	328.81***
# trainees	27.35	4.32	23.03***
# highly educated trainees	6.17	.02	6.16***
% highly educated trainees in total employment	2.64	.01	2.63***
% highly educated trainee hires in total hires	11.7	.18	11.51***
% highly educated trainees workers aged <30years	13.2	.14	13.06***
Inv. per worker (in €1,000)	18.64	21.7	-3.06
	Selected industries		
Manufacturing	.19	.18	.01
Construction	.10	.22	-.12***
Business services	.23	.13	.10***
Public administration	.23	.15	.08***
Number of firms	578	1,725	
Number of observations	3,393	10,023	

Notes: Average values across 1997–2000 of training and non-training firms. Δ: Average in training firm minus average in non-training firms. A training firm is defined as a firm with at least one highly educated trainee in 1997 or 1998, and as non-training firm otherwise. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Training versus non-training firms. 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 1997 or 1998. This divides the sample into 578 training firms and 1,725 non-training firms. Table 2 shows summary statistics for training and non-training firms in the years prior to the reform (1997–2000). Training and non-training firms are fundamentally different. Training firms are four times as large in employment as non-training firms, operate more often in the business service and public administration sector, and less often in construction. Per worker, training and non-training firms invest similar amounts of money.

Each training firm employs on average 6.2 highly educated trainees per year. While highly educated trainees make up only 2.6% of a training firm's workforce, they constitute a large proportion of those workers associated with being able to make a change: 11.7% of a training firm's hires are highly educated trainees, and 13.2% of a training firm's workers below 30 years are highly educated trainees. Common occupations for highly educated trainees are media service occupations, retail occupations, insurance and financial service occupations, or technical drafter, but they also work in manufacturing jobs. Likewise, highly educated trainees are most common in the business service sector, but can also be found in the manufacturing sector.

Investments. Each year, firms in the Establishment Panel are asked whether they invested in four investment types in the last year: (1) production facilities, plant and equipment, furniture and fixtures, (2) communication technology, electronic data processing, (3) real estate and buildings, and (4) means of transport, transportation systems. If a firm invested in at least one of these categories, 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 intangible assets other than ICT are included. Table A1 provides a detailed description of the underlying survey questions.

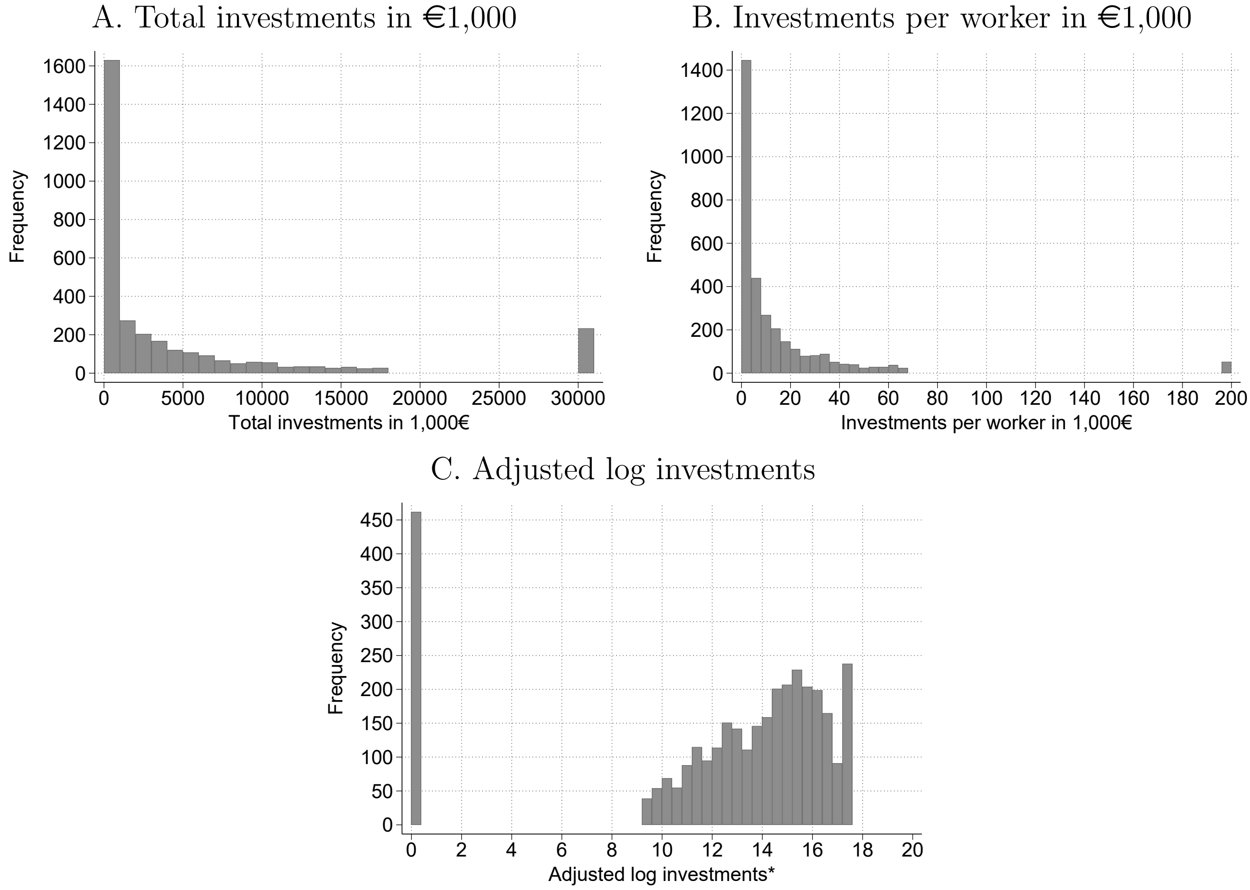
4% of firms never invest throughout 1997–2006, among training firms only 2.3%. Figure 2, Panel A shows the distribution of total investments within the sample of training firms. 13.6 of firm-year observations show no investments, while the distribution of strictly positive investments is highly right-skewed. To curtail the impact of extremely large investments, investments in the highest percentile of the distribution of total investments or investment per worker are trimmed.

I construct two main investment variables: First, I divide total investment by the number of workers to account for the fact that large firms tend to make large investments and to purge the distribution from the right-skewness caused by the right-skewness in firm employment size. The resulting distribution is show in Figure 2, Panel B. The distribution is still right-skewed, with a mean of €23,000 per year per worker, and a median of €8,500.

Second, I employ an adjusted log transformation. Since a simple log-transformation is known to be problematic (e.g. Chen & Roth, 2023), I instead employ an alternative transformation suggested by Chen & Roth (2023): I log-transform investments for strictly positive values and define a change from zero to any strictly positive investment to be as important as an investment increase by 1%, i.e. I manually define $log(0) := -0.01$. Results are robust to that specification. The distribution of this variable is shown in Figure 2, Panel C. I also study the extensive and intensive investment margin separately.

The establishment panel lacks a direct measure of the capital stock. To fill this gap, I exploit information on total investments, the proportion of net investments, dummy variables representing the four investment types, and industry. I apply the modified perpetual inventory method developed by Müller (2008, 2017) explicitly for this dataset to impute the capital stock. I establish a starting value for the capital stock using investments in the first three observed

Figure 2: Distribution of firm investments



Notes: Observations at the firm-year level. Among training firms only. Adjusted log investments: $log(\text{inv})\,\forall\,\text{inv} > 0; log(\text{inv}) := -0.01\,\forall\,\text{inv} = 0$, reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin.

years (1996, 1997 and 1998 at the earliest) and project the capital stock for subsequent years using investment information and sector-specific depreciation rates. Please note that the capital stock is therefore highly unreliable in the first three years and becomes more accurate over time. However, acknowledging the inherent inaccuracies in this method, I focus on investments while reporting results for the log capital stock only in order to assess the effect size.

Technological change. To determine whether investments incorporate new technologies, I use two additional pieces of information from the data: The technical status of a firm's machinery, and firm-level organizational changes, each of which I describe in the following. Unfortunately, information on investment types (production facilities, ICT, real estate, and transport) only distinguishes between "no" or "any" investments in this category. Since the subsequent analyses reveal that the intensive investment margin is much more affected than the extensive investment margin, these variables will not be used to study technology adoption.

Firms are asked to assess the overall technical status of their production equipment compared to other establishments in the same industry on a scale from 1 ("completely out-of-date") to 5 ("state-of-the-art"). Out of all the firm-year observations, 0.5% rate the technical status of their machinery as the lowest category 1, 3% assess it as category 2, 31% as category 3, 49% as category 4, and 16% as the highest category. There is variation in technical status within firms over time: In 35% of the firm-year observations, firms' technical status changes from one year to the next.

Firms also 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 the following four binary indicators: 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-benefit calculations. Most firm-year observations do not include any organizational change (59%), 22% one organizational change, 12% two changes, 6% three changes, and 2% four changes.

4 Event study approach

The identification strategy exploits the quasi-random assignment of the education reform to federal states that produces exogenous variation in the supply of upper-track school graduates across states and years. I compare firms in treated and control states before and after the reform in a difference-in-differences (DiD) event study design by estimating the following specification:

$$Y_{jt} = \sum_{t=1997, t\neq 2000}^{2006} \beta_t(\text{Treat}_j \times \text{Year}_t) + \psi_t + \phi_j + \epsilon_{jt} \tag{1}$$

where Y is one of several outcomes such as investments, j denotes the firm, and t the calendar year.[17] Treat is a binary variable with Treat $= 1$ if the firm is located in a state undergoing the reform and zero otherwise. ψ_t captures calendar-year fixed effects. Firm fixed effects ϕ_j capture time-constant level differences between firms. The vector $\beta_t, t \geq 2001$ includes the coefficients of interest, namely the differential firm outcomes in treated states compared to control states following the reform in 2001 purged from the baseline difference between treated and control firms in 2000. The event study thus identifies the causal effect of a firm facing a state-wide negative trainee supply shock.[18] 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). For brevity, I sometimes use the equivalent difference-in-difference specification, aggregating years in the pre-period 1997–2000, the roll-out period 2001, the post

[17] I stop in 2006 because of a different education reform affecting trainee supply from 2007/2008 onward.

[18] Note that this is different to the causal estimate of a firm employing one fewer trainee.

period 2002–2003, and the phase-out period 2004–2006, and estimating:

$$Y_{jt} = \delta_1(\text{Treat}_j \times \text{Roll-out}_t) + \delta_2(\text{Treat}_j \times \text{Post}_t) + \delta_3(\text{Treat}_j \times \text{Phase-out}_t) + \xi_t + \lambda_j + u_{jt} \tag{2}$$

where the coefficient of interest is δ_2, the difference in the post-period compared to the pre-period for treated compared to control firms.

I estimate equations (1) and (2) for training firms, defined as those firms employing at least one highly educated trainee in 1997 or 1998. The implicit assumption is that treated training firms that were training prior to the reform would have wanted to continue training during the time of the reform. The reform has a direct impact on training firms, while non-training firms are affected only via spill-over effects. I therefore rerun the regression for non-training firms as a falsification test and expect much smaller estimates.[19] Here, the implicit assumption is that treated non-training firms did not want to train in the absence of the reform.

Matching. Treated training firms may differ from control training firms in aspects which expose them to different potentially confounding factors. To ensure that treated and control training firms are comparable, and therefore exposed to similar potential confounders, I match firms based on their pre-treatment characteristics in two steps. First, I match firms within training and non-training status, and nine industry groups. By matching within industries, the estimated reform effects are devoid of distortions arising from industry-specific shocks. Second, I perform a 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, both averaged over the years 1997–2000, and the increase in pre-treatment log employment between 1997 and 2000.[20] To avoid further limiting the size of the sample, I keep the three control firms with the smallest Mahalanobis distance for each treated firm. To ensure good comparability, I subsequently discard the worst 10% of all matches. Results are robust to both aspects. I present findings for both the entire sample of firms and the matched sample throughout the paper. Convincingly, results are similar for both samples.

Table 3 shows characteristics of treated training firms compared to control training firms for both the unmatched and matched sample. Prior to matching, treated firms are significantly smaller and invest less than control firms. The matching works well in eliminating differences in observable firm characteristics, both targeted and non-targeted ones.

The identification of the causal effect via the difference-in-differences event study relies on

[19] Since training and non-training are hardly comparable, and likely interact with each other, I refrain from comparing them directly.

[20] This requires firms to be present both in 1997 and 2000, reducing the extent of panel attrition: 47% of matched training firms present in 1997 are still present in 2006, see again Figure A1.

Table 3: Balancing table

	Unmatched			Matched		
	Treated (1)	Treated - Control (2)	SE (3)	Treated (4)	Treated - Control (5)	SE (6)
			A. Targeted variables			
Δ log employment	-0.26	-0.10	0.07	-0.17	-0.03	0.03
Log employment	4.93	-0.35	0.13**	5.06	-0.17	0.13
Share highly educated trainees	2.86	-1.83	3.52	2.57	0.36	0.36
			B. Non-targeted variables			
# highly educated trainees	3.94	-3.33	1.86*	4.53	-0.60	0.87
Trainee wage	21.75	0.15	0.84	20.03	-0.43	0.55
Adjusted log investments	11.42	-1.15	0.42**	12.65	-0.34	0.48
Inv. per worker in €1,000	17.46	-0.26	3.17	20.82	0.74	3.50
Technical status	3.95	0.05	0.06	3.95	0.04	0.07
Organizational changes	1.15	-0.13	0.11	1.13	0.02	0.11
Number of firms	578			393		

Notes: Δ log employment refers to the change in log employment between 1997 and 2000. All other values are the averages of the values 1997–2000. SE: Standard error. Among training firms only. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

three main assumptions.

Assumption 1 - Parallel trends. First, I assume that firm outcomes in treated states in the 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.

While the matching procedure further improves the plausibility of counterfactual parallel trends, it does not provide remedy if external factors unrelated to the matching variables evolve differently in treated and control states post 2000. As shown in Appendix B, key state metrics such as unemployment, population size, education expenditure, and public debt and investments do not change significantly in treated compared to control states post 2000.[21] Likewise, I argue in Appendix B that the concomitant investment tax programs studied in Lerche (2022) and Siegloch et al. (2024) are unlikely to confound the effect of the education reform. 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

[21]Zooming in on population growth and the unemployment rate, I observe very comparable patterns across states, see Figure C1. 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. Since population size might react to the reform, i.e. inhabitants moving out of the state, I do not focus on the number of 18-years old in 2001 but on the number of 14-years old four years prior to 2001. 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.

Trade Organization in 2001/2002 might confound the reform effect. However, these shocks likely affected treated 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 of this.

Students may have also anticipated the reform. There was, however, very little scope for them to react: When the reform was decided, students of the missing graduation cohort were in grade 7 in Mecklenburg-Western Pomerania and 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. Also, school graduates may delay or accelerate the start of their vocational training in response to the shock. 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 me to identify cross-state commuting. Trainees in the affected states rarely commute across states (2.7% in 1999 to 2001) compared to workers with a university degree (5.3%), 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 mobility rate is extremely low at 2.2%. Further, 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, if trainees moved or commuted from control states to treated states in response to the reform, this would bias the estimates 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 temporary decrease in the supply of trainees. This requires that no other aspect of the reform affects investments.

One other, potentially confounding, aspect of the reform is the increase in the skill level of highly educated trainees due to the increased years of schooling from 2002 onward. Also, the share of upper-track graduates who start vocational training might have been affected. However, these two aspects changed permanently, such that effect dynamics will help distinguish between these permanent adjustments and the temporary trainee shortage.[22]

[22] Also, higher levels of education would, if any, likely induce more investments, and therefore provide a lower

One might be concerned about demand changes accompanying the supply shock. Labor supply changes due to migration usually entail this issue. However, I only study a postponement in the start of vocational training, which is unlikely to affect consumer demand in a meaningful way because, first, the overall population size remains constant, and second, trainees earn low wages, so their demand arguably does not decrease significantly because they are in school one more year. Turning to firm demand, low trainee wages also prevent a meaningful decrease in the firm wage bill when trainee employment is reduced, making it unlikely to present a confounding channel.

Another concomitant aspect of the reform is the potential substitution of missing trainees with workers of a different observed or unobserved type.[23] However, I do not interpret such substitutions as a source of bias but as a mechanism via which the effect unfolds. Besides, I will show empirically that substitutions were very limited.

Trainee distribution across firms. Even if the estimated parameters of interest, $\hat{\beta}_t$, identify the unbiased effect of facing a trainee shortage, they are subject to the realized distribution of trainees across firms. In particular, $\hat{\beta}_t$ are small if trainees are primarily missing in firms that would not have invested in the absence of the shock, and $\hat{\beta}_t$ are large if trainees are primarily missing in firms that would have invested in the absence of the shock. In order to identify the effect on investments independent of the realized distribution of trainees across firms, I propose a complementary identification strategy in Appendix D: I predict the distribution of trainees across firms based on a Bartik-style instrument of firms' pre-reform use of trainees and the state-level shift in trainee employment induced by the reform. This allows me to identify a different causal parameter, namely the effect of employing one trainee fewer. This analysis is, however, more demanding and subject to further assumptions, which is why my preferred identification strategy is the difference-in-difference event study design.

Inference. Standard errors are commonly clustered at the level of treatment assignment to account for cluster-level shocks (e.g. Abadie et al., 2023). Here, this would result in a small number of clusters, i.e. federal states. For valid inference with a small number of clusters, I follow Roth et al. (2023), and assume that any state-specific shock is small compared to the idiosyncratic error terms at the firm level, potentially resulting in a small violation of parallel trends. This assumption is well justified in the data: For the two main outcomes, trainee employment and investments, the variance of the error term is approximately equal to the average variance of the error term within states but much smaller within firms, suggesting little to no within-state correlations but large within-firm correlations, see Table C1. I hence cluster

bound of the effect.

[23] Highly educated 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.

standard errors at the firm level. Additionally, I perform permutation (Fisher randomization) tests, comparing the t-statistic of the treatment effect for the actual treatment assignment and all permuted treatment assignments across federal states. Finally, I also report alternative confidence intervals based wild t-bootstraps clustered at the state level as suggested by Cameron et al. (2008).

5 Bite of the reform

Effect on trainee employment. Figure 3 displays the results of estimating the difference-in-differences event study model outlined in equation (1) regarding the effect of the reform on the employment of highly educated trainees. Endorsing the identifying assumption of parallel trends, firms' highly educated trainee employment evolves in parallel in control and treated states in 1997–2000. In 2002, 2003 and 2004, approximately 1.1 highly educated trainees fewer work in treated training firms compared to control training firms on average. Considering the typical training duration of three years, these are precisely the years the majority of the missing school graduates would have undergone vocational training.[24] The effect is similar in the sample of unmatched and matched firms, and both are statistically different from zero. The effect corresponds to a drop by one-fifth of the average firm employment of highly educated trainees per training firm (see Table C2), and hence less than one-third that would be missing if none of the missing highly educated trainees were replaced. Consistent with the timeline of the shock, the employment gap starts to shrink for the matched sample in 2005. However, the gap does not close immediately indicating delays, for example, related to military service.

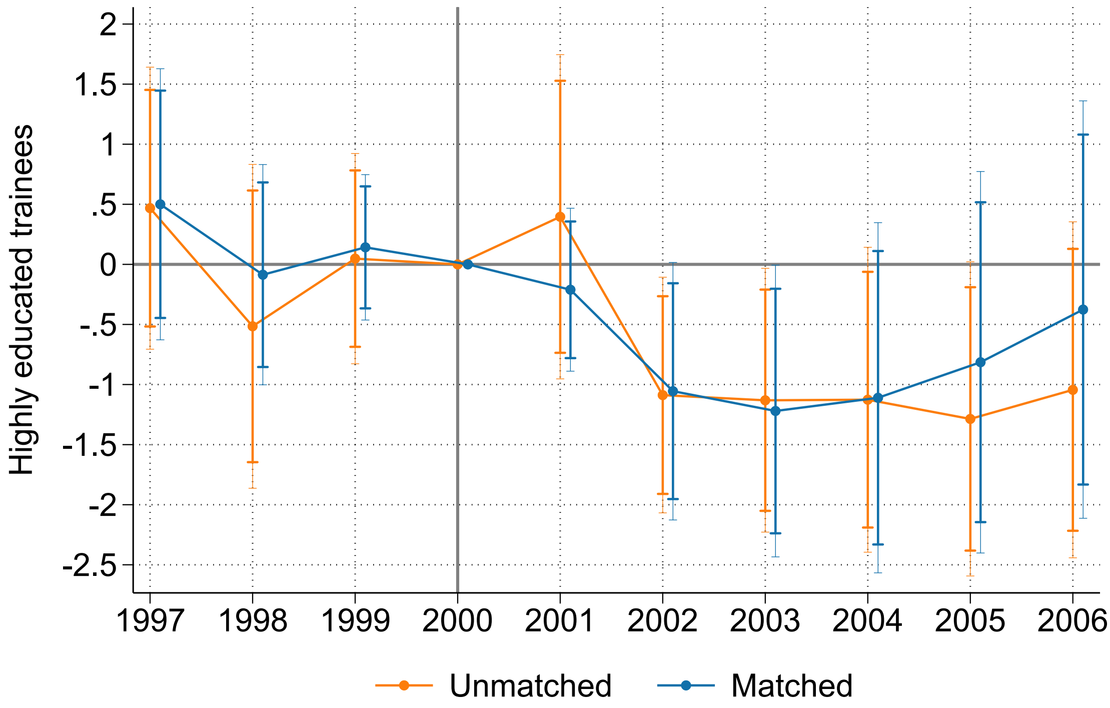
Wage effects. The detailed administrative labor market data allows me to study firms' adaptation strategies, such as changes in trainee wages, or the substitution of highly educated trainees with other workers. To investigate such effects, I employ the corresponding difference-in-differences specification given in equation (2), comparing the pre-treatment period 1997–2000 to the post-treatment period 2002–2004. Results are given in Table 4.

There is no evidence of an increase in the wages of highly educated trainees in response to the negative supply shock (column 1). This is in contrast to what standard economic theory predicts. To understand the absence of any wage effects, it is important to keep in mind that trainee wages are very rigid, often set by collective bargaining agreements, that the shock was only temporary, and that the supply of highly educated school graduates is fixed by the cohort size, giving very little scope for wage increases to increase their employment.[25]

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

[25]In fact, there are a multitude of reasons that potentially explain the lack of a wage adjustment. First, firms likely shy away from increasing wages in response to a temporary shock because downward rigid wages will impede a subsequent wage decline once the supply shock dissipates. Second, trainee wages in Germany are set at a very low level and are paid only throughout the three-year vocational training period. Hence, even a

Figure 3: Effect on employment of highly educated trainees



Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (1). Standard errors clustered at the firm level. Among training firms only. For the corresponding graph with confidence intervals based on cluster wild t-bootstraps, see Figure C2. N=3,322 for the unmatched sample and N=3,182 for the matched sample.

Worker substitution effects. Prominent candidates acting as substitutes for the missing highly educated trainees are low-educated trainees, highly educated trainees from other federal states, and highly educated workers who have already completed vocational training. However, firms do not compensate for their missing highly educated trainees by hiring more low-educated trainees (column 2). In consequence, overall trainee hires also drop. The low substitutability between low- and highly educated trainees, in line with Muehlemann et al. (2022), is likely related to distinct skill sets, the specialization in different occupations, and the unchanged demand for low-educated trainees against a fixed supply of school graduates.

Also, there is no statistically significant increase in cross-state commuting of highly educated trainees from a different federal state following the shock (column 3), even though the coefficient is positive. No increased commuting supports the SUTVA assumption of no spill-overs across state borders.[26]

hypothetical doubling of training wages would result in negligible changes in absolute lifetime income. Instead, trainee supply responds to anticipated post-training wages (Neuber-Pohl et al., 2023) that remain unchanged in the present case. Third, the vast majority of training wages are set by collective bargaining agreements, and even firms that are not part of those agreements tend to base their wages on such agreements. Of course, firms could deviate upwards. In that case, worker's councils, which would have to approve training wages in large firms, would likely oppose unequal treatment of trainees. Finally, this finding is in line with the results by Muehlemann et al. (2022) in the case of the opposite, positive supply shock of trainees.

[26]The coefficient of interest captures potential increases in commuting into treated states in addition to

Table 4: DiD Results – Wage and worker substitution effects

	Log wage highly educ. trainees (1)	# low-educ. trainees (2)	# highly educ. commuting trainees (3)	Log highly educ. VT employment (4)	Log wages educ. VT employment (5)	Trainee retention rate (6)	Internal retraining (7)
			A. All training firms (Unmatched)				
Treat × Post	-0.00	-0.90	2.46	0.01	-0.02	-0.06	-0.27*
	(0.03)	(1.77)	(3.18)	(0.08)	(0.02)	(0.04)	(0.16)
Mean dep. variable	3.02	10.06	2.83	2.18	4.29	0.65	0.47
N	2252	3322	1429	3083	3082	3150	1618
			B. Matched training firms				
Treat × Post	0.01	-1.65	2.04	0.03	-0.02	-0.09**	-0.09
	(0.03)	(1.76)	(2.87)	(0.09)	(0.03)	(0.04)	(0.07)
Mean dep. variable	3.03	9.72	2.93	2.25	4.31	0.64	0.50
N	2198	3182	1564	3032	3031	3035	1586

Notes: Difference-in-difference coefficients based on equation (2). Pre: 1997–2000. Roll-out: 2001. Post: 2002–2004. Fade-out: 2005–2006. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Among training firms only. Mean dep. variable: Average outcome of treated firms in 2000. Column 3: a commuter is defined as a person living and working in two distinct federal states. This variable is available from 1999 onward. Column 6: 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. Column 7: Internal retraining is the sum of retraining incidences at the firm-year level. VT: completed vocational training. For the full set of results, see Table C3.

Columns 4 and 5 show that the employment and wages of highly educated workers who have completed their training program do not increase in response to the trainee shortage, indicating that already trained workers are not suitable substitutes for trainees.[27]

Firms may try to compensate for missing trainees by retaining more trainees upon training graduation. Likewise, poaching of these workers might increase as well. I find that the retention rate of recently graduated trainees decreases (column 6) in response to the shock, suggesting increased poaching.

Firms may also increase retraining of incumbent workers to overcome skill shortages caused by the negative trainee supply shock. In contrast, I observe a decline in internal training measures in treated training firms by approximately one-third of the initial value (column 7). This finding may be related to foregone technology adoption and foregone organizational change, as I show below.

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

potentially reduced commuting into control states, and thus provides an upper bound of the true effect.

[27]Likewise, I find no evidence of substitution with low-educated workers who have completed their training program, and the subgroup of highly educated workers who have completed their training program below the age of 30.

6 Effects on firm technology investments

6.1 Effect on investments

I now turn to the reform effects on firm investments, interpreting them as effects of the negative trainee supply shock. Figure 4, Panel A, shows the difference in investments per worker between treated training firms and control training firms over time. Convincingly, there are no statistically significant pre-trends, supporting the assumption that investments in treated states would have evolved in parallel with investments in control states in the absence of the trainee supply shock. I find a statistically significant decline in investments per worker following the reform in treated training firms compared to control firms; the key finding of this paper. The pattern looks comparable among the unmatched and matched sample of firms, and when using a different definition of the investment variable, namely adjusted log investments (see Panel B),[28] though the estimates are less precise.

The decrease in investments is temporary: investments of treated firms catch up with investments of control firms in 2004/2005,[29] but do not overshoot. The temporary drop in investments when trainee employment is temporarily reduced, in combination with a lack of anticipation and overshooting, suggests that investments and trainees are complements.

The investment decrease is found among small and large firms, as well as among firms in the skilled business service and public administration sector—where capital and labor tend to be complements—, in the manufacturing sector—where capital and labor tend to be substitutes—, see Table C5.

The estimated average decline in investments is large given that highly educated trainees make up only 2.6% of a training firm's workforce: Investments are reduced by approximately €6,200 and €5,300 per worker in 2002 and 2003 (matched sample; €5,300 and €4,000 for the unmatched sample), or 0.7 and 1.1 log points (matched sample; 0.6 and 1.2 for the unmatched sample). Hence, the investment decline goes beyond a potential "mechanical" effect of reducing capital in proportion to trainee employment.[30] The effect corresponds to a decline of 12–24% of a standard deviation of firms' investments (20–40% of the within-firm standard deviation), see Table C2. Please note again that these numbers represent the sample and not the population average treatment effect.

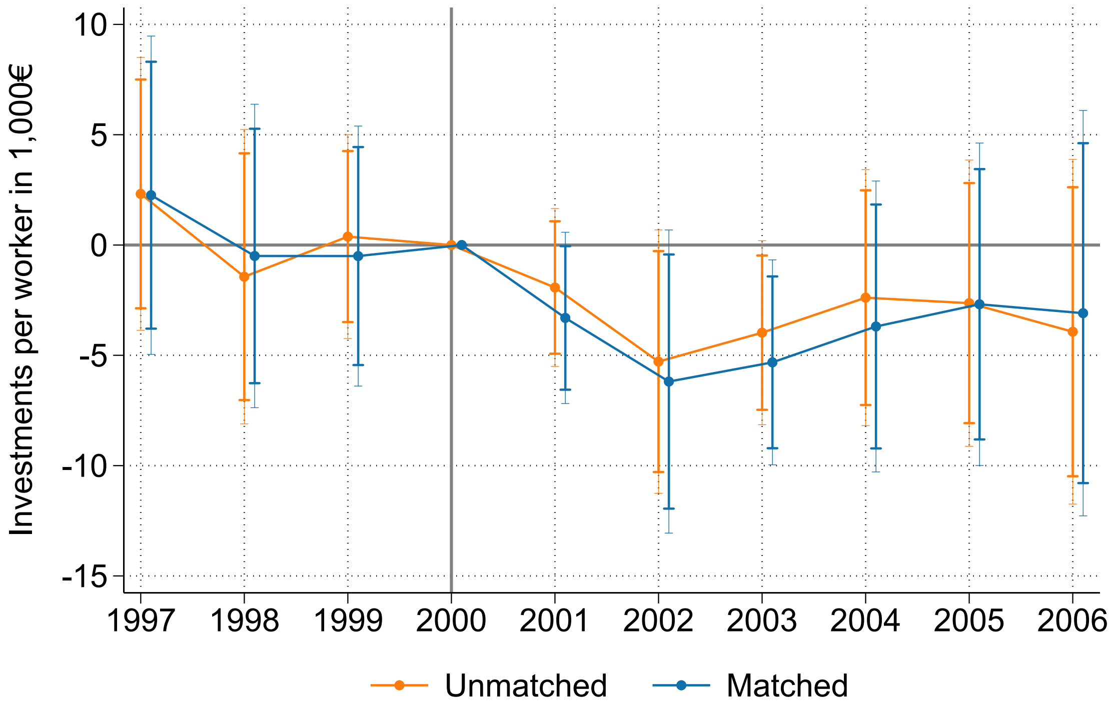
[28] Log(investments) for strictly positive investment values. For zero investments, the outcome variable is set to -0.1, reflecting that a change at the extensive margin is valued ad much as a 1% change at the intensive margin.

[29] This is why in the subsequent difference-in-differences analyses I will use the years 2002–2003 as post period.

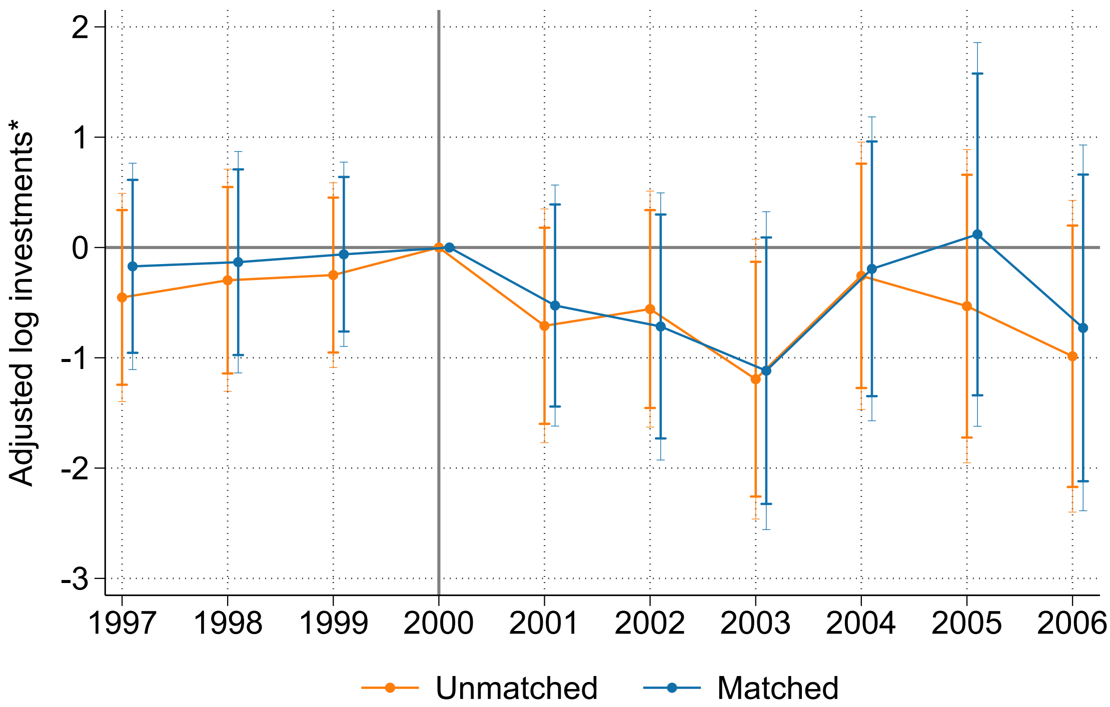
[30] Likewise, the representative BIBB-Cost-Benefit-Survey 2000 suggests that the "mechanical" costs are much smaller than the estimated effect: East German firms surveyed in 2000 spent an average of €487 per trainee per year on equipment and material (Beicht et al., 2004). With a reform-induced reduction in the number of trainees of 1.50 in 2002 and an average training firm size of 354 workers this would imply a "mechanical" reduction of €2.06 per worker. In addition to the €487 spent on equipment and material costs, in 2000 East German firms reported €1,530 of "other costs" per trainee per year, including costs for teaching material, fees, and training administration. If a firm interpreted all of these costs as capital investments, the total "mechanical" reduction in investments would still be as small as €8.55 per worker.

Figure 4: Effect on investments

A. Investments per worker in €1,000



B. Adjusted log investments



Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (1). Standard errors clustered at the firm level. Among training firms only. Outcome Panel A: investments in €1,000 divided by total employment in 1997. Outcome Panel B: Adjusted log investments: $log(\text{inv})\,\forall\,\text{inv} > 0; log(\text{inv}) := -0.01\,\forall\,\text{inv} = 0$, reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin. For the corresponding graph with cluster wild t-bootstrap confidence intervals, see Figure C3. N=3,322 for the unmatched sample. N=3,182 for the matched sample.

To get a better understanding of the magnitude of the effect, it is useful to consider the effect on the log capital stock even though the capital stock is only an imperfect imputation based on investments and assumed depreciation rates (see again Section 3). Estimates based on the difference-in-differences specification suggest that the capital stock decreases by 7–10%, see Table 5, column (1).

Comparing the effect with estimates from the literature is problematic because few other shocks are temporary. The perhaps most appropriate comparison is based on a temporary (three-year-long) bonus depreciation in capital costs in the US between 2001 and 2010, evaluated by (Zwick & Mahon, 2017), who find a price elasticity of 7.2. Using this estimate, the decrease in investments in response to the trainee shortage corresponds approximately to the decline if capital costs increased by 7.3–16.6%.[31]

Extensive versus intensive margin. The estimated effect can be decomposed into an extensive and intensive investment margin effect, as Table 5, columns 2 and 3 show based on the corresponding difference-in-differences estimation. The results reveal that the extensive margin—measured as the binary outcome of investing versus not investing—is unaffected, while the intensive margin—measured as log investments—adjusts, even though the estimates lack statistical significance. This implies that firms forego large investments in response to the trainee shortage. To explicitly test this hypothesis, I run the difference-in-differences regression among observations with strictly positive investments using a binary outcome taking the value one for investments in the upper tercile of the investment per worker distribution (>€10,000), and zero otherwise, see Table 5, column 4. Treated training firms are 11–16 percentage points less likely to make large investments than control training firms when trainees are scarce.[32] This finding is in line with the idea that investments are usually both costly and indivisible, said "lumpy" (e.g. Cooper et al., 1999; Bessen et al., 2020). At the same time, not all firms are constantly exposed to adopting new technologies. Therefore, firms do not optimize investments over a continuous investment distribution but face a discrete investment choice. In the setting of this paper, this implies that some firms, not planning to invest regardless, do not reduce investments. However, others, intending to make large, lumpy investments, forego these plans due to the trainee shortage.

[31]The smallest treatment effect (matched firms, 2001) of -0.53 corresponds to a the decline if capital costs increased by 0.526/0.072=7.3. The largest treatment effect (unmatched firms, 2003) of -1.194 corresponds to a the decline if capital costs increased by 1.194/0.072=16.6.

[32]The effect is comparable when focusing on investments per worker in the upper decile of the distribution, and even more pronounced when defining large investments within industries.

Table 5: DiD Results – Additional investment effects

	Log(K) (1)	Any inv. (0/1) (2)	Log(Inv.) (3)	Large inv. (1/0) (4)
	A. All training firms (Unmatched)			
Treat × Post	-0.07	-0.02	-0.16	-0.11**
	(0.05)	(0.04)	(0.15)	(0.05)
Mean dep. variable	10.18	0.90	13.98	0.33
N	3155	3308	2843	2843
	B. Matched training firms			
Treat × Post	-0.10*	-0.03	-0.24	-0.16***
	(0.06)	(0.04)	(0.16)	(0.05)
Mean dep. variable	10.04	0.89	13.82	0.30
N	3064	3176	2809	2809

Notes: Difference-in-difference coefficients based on equation (2). Pre: 1997–2000. Roll-out: 2001. Post: 2002–2003. Fade-out: 2004–2006. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Among training firms only. Mean dep. variable: Average outcome of treated firms in 2000. Log(K): Log of the imputed capital stock. Large inv.: Investments in the upper tercile of the distribution of strictly positive investments per worker assigned as one, and zero otherwise. For the full set of results, see Table C4.

Robustness. The negative effect of the reform on investments is not only robust to the definition of the outcome variable—investments per worker or adjusted log investments—, and to matching in treated and control states, but also to a wide range of further specifications. To see this, I present the difference-in-differences event study estimates for both investment variables for the year 2003 based on the matched firm sample in Figure 5.

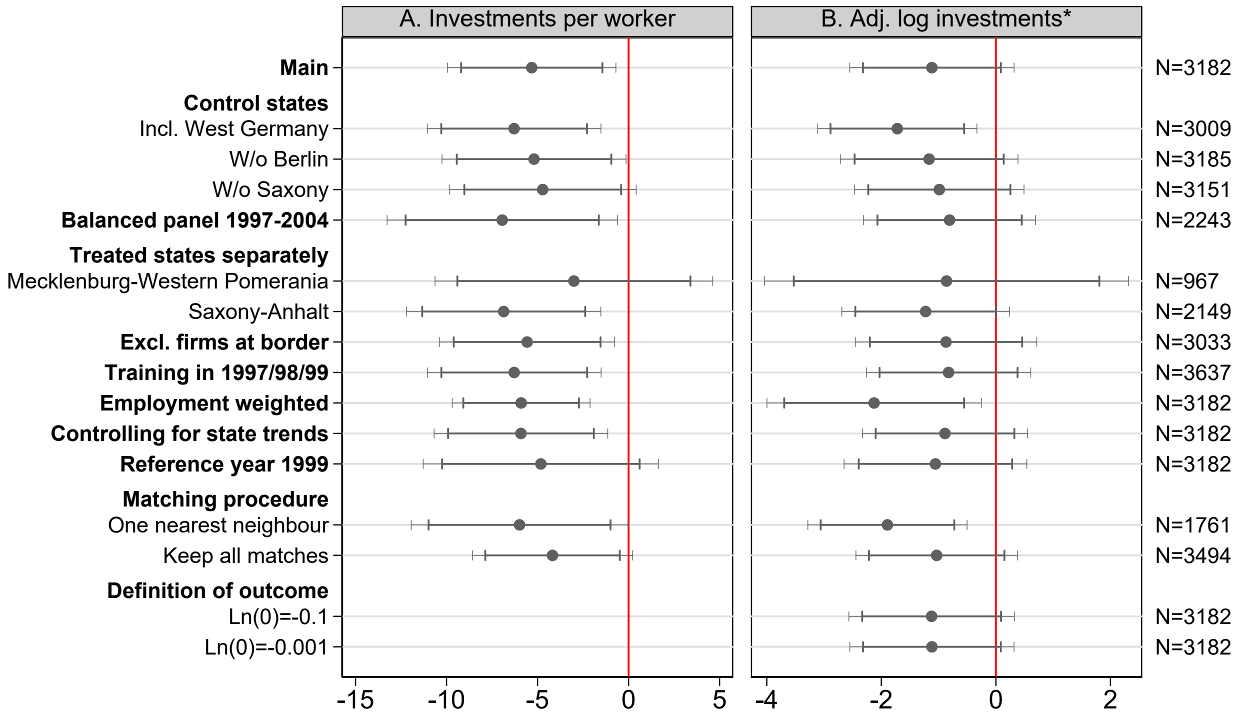
First, I expand the set of control firms to include West German firms which were initially discarded because they may arguably be substantially different to East German firms as well as exposed to different shocks. Including them in the set of control firms produces significantly negative estimates that are slightly larger in magnitude. When excluding Berlin or Saxony-Anhalt from the set of control states, due to their slightly different demographic and economic trends, the results also remain robust.

Firms in the main sample may disappear over time. The negative estimate is robust to a restriction of the sample to firms observed for each year between 1998 and 2004, but is less precisely estimated.

Convincingly, the effect is found within both treated states, despite their differences in industry structure and geography. However, the effect is significantly smaller and less precisely estimated for firms in Mecklenburg-Western Pomerania, which is based on a very small sample of firms, than for firms in Saxony-Anhalt.

Firms at federal state borders might be less affected by the reforms because they may attract trainees from control states. Since firms' addresses are not disclosed in the data, and counties are too large to reliably identify firms close to the federal state border, I instead use the share of commuters across federal states as a proxy for worker supply from other states. Excluding firms with a commuter share in the highest decile in 1999 does not meaningfully affect the results.

Figure 5: Robustness



Notes: Event study coefficients and 90% and 95% confidence bands of the term Treat × 2003. Panel A using investments per worker in €1,000 as outcome; Panel B using adjusted log investments as outcome. Adjusted log investments: $log(\text{inv})\,\forall\,\text{inv} > 0; log(\text{inv}) := -0.01\,\forall\,\text{inv} = 0$, reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin. Standard errors clustered at the firm level. Among training firms only. N indicates the number of observations in the respective estimation. *Main:* Main specification. *Control states:* Additionally including all West German training firms as control firms, or dropping Berlin or Saxony from the set of control firms. *Balanced panel 1997-2004:* Sample restricted to firms observed in each year between 1997 and 2004. *Treated states separately:* Only using treated firms from one treated state and dropping firms from the other treated state. *Excl. firms at border:* Dropping those 10% of firms with the highest 1999 cross-state commuter share of workers with vocational training. *Training in 1997/98/99:* Training firms defined as those with at least one highly educated trainee in 1997, 1998, or 1999 instead of 1997 and 1998 only. *Employment weighted:* Observations weighted by firms' initial employment size in 1997. *Controlling for state trends:* Additionally controlling for linear state-specific time trends. *Reference year 1999:* Using 1999 instead of 2000 as reference year. *Matching procedure:* Using only the nearest neighbor instead of the three nearest neighbors as control firms, and keeping all matches instead of discarding the furthest 10% of all matches. *Definition of outcome:* Assigning $log(0) := -0.1$, and $log(0) := -0.001$ instead of $log(0) := -0.01$.

Next, I alter the definition of training firms which was initially based on the years 1997 and 1998 to minimize potential anticipation. Defining a firm as training firm if at least one highly educated trainee was employed in either 1997, 1998 or 1999 does not meaningfully affect the results.

When weighting the observations by the firms' initial employment size in 1997, the coefficient remains negative.

Acknowledging that states may be on different (linear) time trends, and controlling for them, does not meaningfully affect the results. Results are also robust to using 1999 instead of 2000 as the reference year, though precision declines.

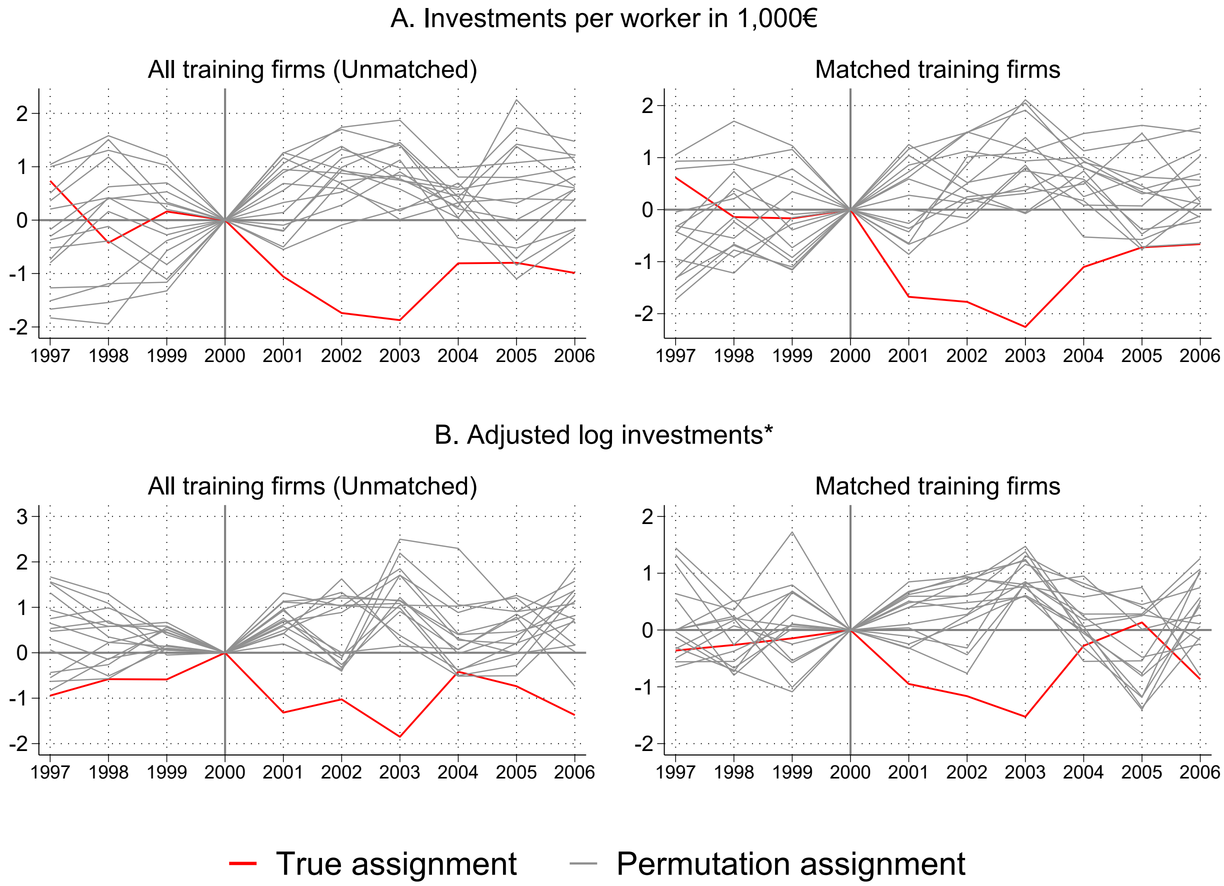
The results are also robust to different specifications of the matching procedure. When using the nearest neighbor instead of the three nearest neighbors, the estimated coefficient remains very similar in Panel A (investments per worker), but even increases in size and precision in Panel B (adjusted log investments), despite a significant reduction in sample size. The results are similarly robust to the inclusion of the 10% most distant matches.

Finally, I investigate whether choosing to set $ln(0)$ to -0.01 (reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin) drives the results observed thus far. Convincingly, the results are virtually unchanged when choosing the values -0.1 and -0.001 (reflecting that a change at the extensive margin is valued as much as a 10% change at the intensive margin, or a 0.1% change respectively).

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 as a valid method for inference when the number of clusters is small (e.g. Roth et al., 2023). Figure 6 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 errors of firms within states. Following the 2001 reform, all panels, i.e. investments per worker, adjusted log investments, and for both the unmatched and matched sample, 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 it is very unlikely that only cluster-level shocks would have caused the observed investment decline. Likewise, no comparable decrease in the employment of highly educated trainees was observed under any permutation assignment, see Figure 6, Panel A. This result holds for both the sample of all firms and the sample of matched firms.

Since the number of possible permutations within East Germany is limited to 15, I repeat the permutation test across the 10 West German federal states. There was no comparable education reform in West Germany around that time. The t-statistics of the highestand lowest 2.5% (5%) of the draws under permuted treatment assignment are shown in Figure C6, Panel B. Again, the t-statistic of actual treatment assignment stands out as an outlier and is much smaller than the 5% and 2.5% most negative t-statistics under permuted treatment assignment. See Figure C7 and C8 for the corresponding results on the effect of trainee employment.

Figure 6: Permutation test – T-statistics



Notes: T-statistics of the event study regression based on equation (1) using the actual treatment assignment (red line) and all possible permutation assignments (gray lines).

Falsification test among non-training firms. To validate that the investment decline is indeed linked to the trainee shortage and not due to some idiosyncratic factor happening in the two treated states around this time, I turn to the sample of non-training firms, i.e. firms with no highly educated trainees in either 1997 or 1998. Assuming that non-training firms voluntarily abstain from training highly educated trainees, and were not forced to do so because of the negative trainee supply shock, non-training firms should be much less affected by the reform, at most via spillovers. Since the sample of non-training firms is much more heterogeneous than the sample of training firms, I match treated non-training firms with control non-training firms following the same procedure as for training firms, except that I do not include the share of highly educated trainees in the Mahalanobis matching procedure. Table 6 shows the event study estimates for both training and non-training firms, considering investments per worker and adjusted log investments as outcomes. While the investment drop among non-training firms is not exactly zero, likely reflecting spillover effects, we see much larger declines in investment among training than among non-training firms. This provides further evidence that the investment drop is indeed related to the negative trainee supply shock.

Table 6: Effect on investments – Training versus non-training firms (Matched)

	Adj. log investments		Inv. per worker	
	Training (1)	Non-training (2)	Training (3)	Non-training (4)
Treat × Post	-0.81	-0.25	-6.11*	-2.11
	(0.57)	(0.33)	(3.13)	(1.33)
Mean dep. variable	12.28	8.75	15.81	9.79
N	3322	9791	3322	9791

Notes: Difference-in-difference coefficients based on equation (2). Pre: 1997–2000. Roll-out: 2001. Post: 2002–2003. Fade-out: 2004–2006. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mean dep. variable: Average outcome of treated firms in 2000. Adjusted log investments: $log(\text{inv})\,\forall\,\text{inv} > 0; log(\text{inv}) := -0.01\,\forall\,\text{inv} = 0$, reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin.

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

The analysis reveals that more exposed training firms indeed experience larger employment decreases of highly educated trainees. Likewise, training firms with larger predicted employment decreases of highly educated trainees reduce investments more. In particular, each missing highly educated trainee reduces firm investments by approximately €550,000, corresponding to 9.4% of yearly average investments in training firms in 2000. 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 suggests that firms reduce investments in times of the trainee shortage even when they manage to employ trainees nonetheless, with this investment reduction either due to spill-over effects across firms within treated states, or due to the negative selection of trainees available during the trainee shortage.

Firm size effect versus investment intensity effect. A supply reduction of young labor market entrants may decrease firm investments because young labor market entrants drive firm investments, or because the supply reduction of young labor market entrants impedes firm employment growth. Indeed, log total firm employment decreases in treated training firms compared to their matched control training firms by 7–12% in 2003–2006, see Figure C4, Panel A. There is no decrease in firm size in the unmatched sample. To understand the importance of the firm size effect, i.e. the decrease in firm size holding investments per worker constant, compared to the investment intensity effect, i.e. the decrease in investments per worker holding firm size constant, I decompose the overall investment effect into the size and the intensity effect:

$$\Delta\Delta\text{LogInv} = \underbrace{\Delta\Delta\text{LogN}}_{\text{Size effect}} + \underbrace{\Delta\Delta\text{Log}\left(\frac{\text{Inv}}{\text{N}}\right)}_{\text{Intensity effect}} \tag{3}$$

with $\Delta\Delta$ the difference in a variable between treated and control firms in a certain year compared to 2000, $LogInv$ denoting adjusted log investments, $LogN$ denoting log total firm employment, and $Log\frac{Inv}{N}$ denoting adjusted log investments per worker.[33] Running the same event study regression as above but with log employment and adjusted log investments per worker as outcomes, see Figure C4, and plugging in the regression estimates in equation (3) yields the numbers shown in Table 7. For the unmatched firm sample, Panel A, the entire investment drop is caused by a decline in investment intensity. For the matched sample, Panel B, the investment decline is initially mostly caused by the investment intensity effect (91% in 2002), while in 2004 the firm size effect explains 39% of the investment decline.

Table 7: Relative importance of the investment intensity effect

	2001	2002	2003	2004
All training firms (Unmatched)	108%	111%	101%	107%
Matched training firms	102%	91%	88%	61%

Notes: Relative importance of the investment intensity effect as opposed to the firm size effect for the total investment effect based on the decomposition in equation (3) and the according event study coefficients.

To sum up, supply reductions of young labor market entrants affect firm investment via both margins, via reducing investments independent of firm size, and via depressing firm employment, with the former mechanism being quantitatively more important. The mechanism in Section 7 therefore focuses on the latter only.

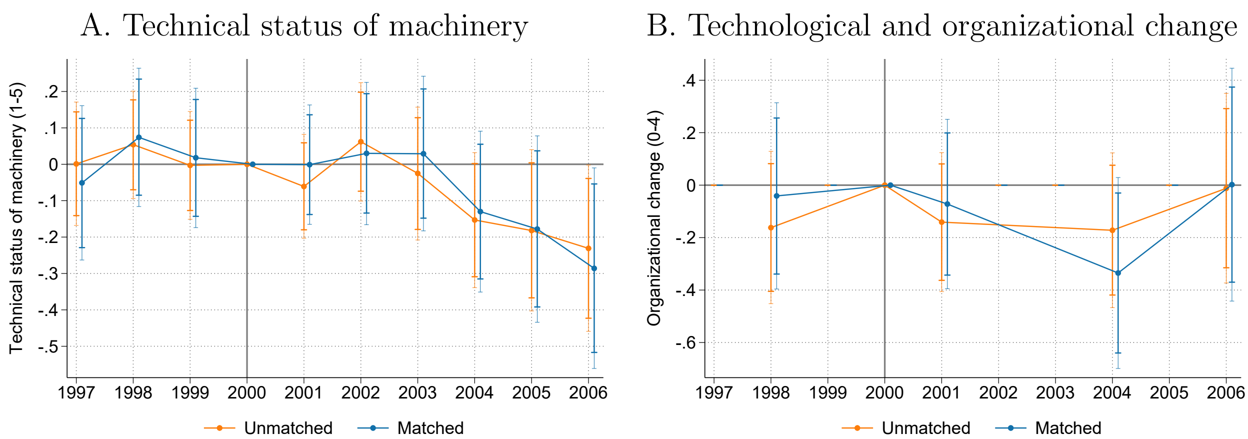
[33] For this decomposition using adjusted log investments as opposed to investments per worker is straightforward as it allows me to rewrite the left hand side as a simple sum. Adjusted log investments per worker are defined equivalent to adjusted log investments as $log(\frac{0}{N}) := -0.01 - log(N)$.

6.2 Effect on firm technology adoption

Having established that the reform-induced trainee shortage decreases overall investments, the following section investigates whether this decrease is linked to foregone technology adoption. The finding that the investment decline is predominately driven by a reduction in the investment intensity, and not by a decrease in firm size, already tells us that the investment drop is unlikely to be (exclusively) driven expansion investments. I now explicitly study the treatment effect on direct indicators of firm-level technology adoption.

In particular, I look at the self-assessed technical status of a firm's machinery on a scale from 1 ('completely out-of-date.') to 5 ('state-of-the-art'). Unlike investments, technical status is a *stock* variable, expected to deteriorate as foregone investments accumulate. I therefore expect the technical status to deteriorate once the missing investments of the years 2001–2004 accumulate. As shown in Figure 7, Panel A, treated training firms report an outdated technical status of their machinery compared to control firms from 2004 onward, with the estimate being statistically significant for both the matched and unmatched sample in 2006. The depreciation is meaningful in magnitude: a depreciation by -0.29 in 2006 (matched, -0.23 unmatched) corresponds to more than half of the within-firm standard deviation, see again Table C2.

Figure 7: Effect on technology adoption



Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (1). Standard errors clustered at the firm level. Among training firms only. For the corresponding graph with confidence intervals based on cluster wild t-bootstraps, see Figure C5.

I also study firm-level organizational change, see Figure 7, Panel B.[34] This approach recognizes that organizational changes often accompany changes in technology, 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 compared to their matched control firms in 2004 of 0.34 reorganization measures less per firm, compared to a mean number of 1.02 reorganization measures and a within-firm standard deviation of 0.75,

[34] This variables is only filled for the years 1998, 2000, 2001, 2004, and 2007.

see again Table C2. Foregone technological and organizational change, in turn, may explain the reduction in internal retraining of incumbent workers established above.

I conclude that at least part of the investment decline is the result of reduced technology adoption. Hence, young labor market entrants are important complements to firms technology adoption. At the same time, it is unlikely that new technologies are complementary to (all) other workers because labor, in general, was not scarce during the trainee shortage.

Foregone technology adoption should affect firm performance in the long-run. However, panel attrition and a confounding trainee supply shock starting in 2007/2008 impede studying longer-term outcomes.

7 Economic mechanism

7.1 Stylized economic framework

I next propose a stylized economic framework that is able to rationalize the established complementarity between young labor market entrants and new technologies. A more detailed formalization is available in Appendix E. Expanding the endogenous technological change model in Acemoglu (1998), I introduce technology vintages requiring vintage-specific skills, and capital adjustment costs of worker training in these new skills, which endogenously make trainees complements to new technologies.

Consider the following set-up within the task framework à la Acemoglu & Autor (2011): Firms maximize profits by deciding whether to adopt a new, exogenously arriving technology. A new technology may substitute or complement labor in existing tasks. Crucially, the new technology always introduces at least one new task.[35] This new task is not a priori assigned to a certain type of worker, e.g. low-skilled or high-skilled, but requires new skills specific to the technology vintage. Consequently, firms incur capital adjustment costs of training workers in this new skill, and the technology complements the worker type that can acquire the new skills at the lowest cost.

Firms can acquire skills either by retraining incumbent workers or by training young, initially unskilled labor market entrants.[36] 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. Consequently, the opportunity costs of training young labor market entrants are lower than the opportunity costs of training incumbent workers, and productivity gains from training young

[35] The literature provides many examples of how new technologies require new skills, without ruling out the replacement of labor in existing tasks (e.g. Autor et al., 2003; Acemoglu & Restrepo, 2018; Deming & Noray, 2020; Autor et al., 2024).

[36] In principle, firms could also acquire these skills by poaching workers who have already acquired the new skills. This, however can never be a stable equilibrium. Also, it comes with other disadvantages for the firms, such as having to invest in firm-specific skills, higher hiring costs, and increased risk when it comes to personnel decisions due to less opportunities for screening. For simplicity, I discard this option.

labor market entrants are higher than the productivity gains from training incumbents.[37] Based on these considerations, firms choose to "make" trainees complements with new technologies.[38] Note that this prediction is independent of whether only the opportunity costs channel, the productivity gains channel, or both apply.

Now, a variation of the standard endogenous technological change argument following Acemoglu (1998) applies: Because young labor market entrants and new technologies are complements, firms adopt fewer new technologies when the supply of young labor market entrants declines. In particular, if young labor market entrants are unavailable, firms adopt a new technology only if the productivity gain from retraining incumbents is large enough to offset the costs of retraining incumbent workers. If retraining incumbent workers is too costly compared to its payoff, technologies which would have been adopted if trainees were present, are not adopted.

This endogenous technological change model with endogenous assignment of trainees to new technologies implies a number of additional hypotheses: First, if the need for new skills is one underlying source behind the complementarity of young labor market entrants and new technologies, the reform-induced investment drop should be larger in firms exposed to strong skill change. Second, if firms "make" trainees endogenous to new technologies out of cost-benefit considerations, firms should be aware of this complementarity. Third, independent of the underlying mechanism, we should observe that young workers work more frequently with new technologies than older workers. These hypotheses are taken to the data in the next section.

7.2 Supporting empirical evidence

The reform-induced investment drop relates to the skill dimension. If the necessity of vintage-specific technology skills is the reason underlying firms' investment reductions, firms that are more exposed to skill change should cut investments to a greater extent in response to the negative trainee supply shock than firms that are less exposed to new skills. Intuitively, firms with incumbent workers in occupations that have not changed recently do not rely on young labor market entrants to invest in technologies because the incumbent workers are still appropriately skilled. In contrast, firms with incumbents in occupations with recent skill change 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 change using changes in vocational training curricula from Lipowski et al. (2024). Changes in vocational training

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

[38] An alternative is that firms either fully adopt a new technology, with all workers using it, or not adopt at all. In this case, all workers are (not) trained the new skills, making them all complements to the new technology. In this scenario, the full transition to the new technology is less costly for the firm the higher the share of entrants in overall employment.

curricula offer an ideal approximation of skill change for three reasons. First, they directly apply to the worker group in question, i.e. trainees and workers who have completed their training program. Second, their changes are often related to technological innovation (Lipowski et al., 2024). Third, they are exogenous to individual firms since they are decided upon at the national level. I approximate firm exposure to new skills as the 1997–1999 average share of workers in occupations whose training curricula were updated around the time of the reform.[39] There is substantial variation in firm exposure to new skills ranging from 0% to 100% with a mean of 68% of a firm's workforce in changing occupations. I standardize the variable to have mean zero and standard deviation of one.

I rerun the event study regression from equation (1), including the triple interaction terms Treat × Year × NewSkills plus all corresponding two-way and one-way interaction effects. Since the goal is to compare investment drops between two treated firms operating in the same industry and with the same exposure to the reform, but with different exposures to new skills, I run the regression in the matched firm sample, assigning the matched control firms the same value of exposure to new skills as the respective treated firm.

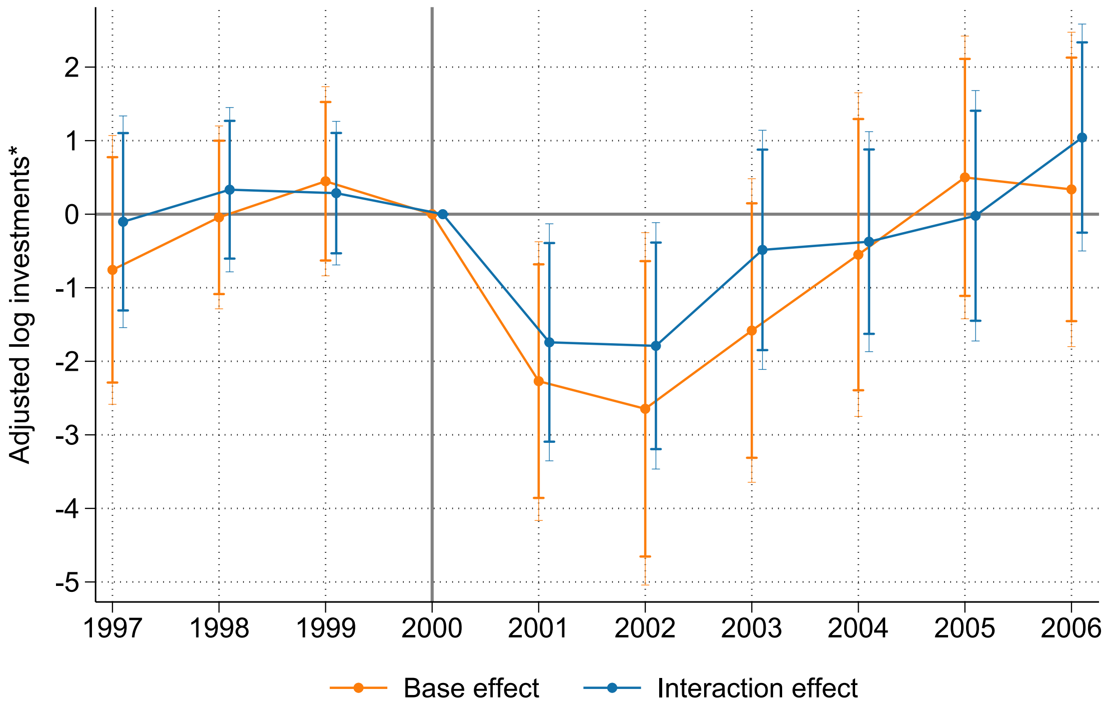
The results are shown in Figure 8. In line with the hypothesized mechanism, the predicted investment drop is larger among firms with stronger skill change. This result is robust to controlling for the triple interaction terms with industry, initial use of highly educated trainees, and firm size. To sum up, there is support in the data for the hypotheses that the need for new skills is one underlying source behind the complementarity of young labor market entrants and new technologies.

Firms acknowledge their need for trainees to adapt to technological change. The framework implies that firms consciously rely on young labor market entrants/vocational trainees to adapt to technological change. This assumption can be directly assessed in representative firm survey data from the BIBB-Cost-Benefit survey. Among all East German training firms surveyed in 2000, approximately half state they use vocational training to ensure a constant supply of new skills and knowledge, while only 16% state they do not do so, see Table 8. Similar numbers are found for using vocational training to improve the firm's adaptability to technical change, and for using vocational training to enhance the firm's innovative capabilities. For more details, see Appendix F.

Young workers use new technologies more often. If firms use young labor market entrants to adopt new technologies, this should be reflected in young workers using new technologies relatively more often than older workers. I test this hypothesis based on a large, representative employee survey in Germany in 1999, 2006 and 2012 that questions respondents

[39]In the main specification I use 103 curriculum changes between 1998 and 2003, for example in the training occupation industrial clerk, metal worker, retail salesmen, or vehicle builder. Results are independent of the exact time period of curriculum changes considered.

Figure 8: Heterogeneity by exposure to vintage skills



Notes: Event study coefficients of the interaction terms Treat × Year and Treat x Year x NewSkills plus 90% and 95% confidence bands Standard errors clustered at the firm level. Among training firms only. Adjusted log investments: $log(\text{inv})\,\forall\,\text{inv} > 0; log(\text{inv}) := -0.01\,\forall\,\text{inv} = 0$, reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin. New skills defined as the 1997–1999 average share of workers in occupations whose training curricula are updated between 1998 and 2003. N=2,846.

in Germany about their main working tools (IAB/BIBB/BAuA Qualification and Career Survey). In particular, I regress the usage of new technologies among workers who have completed their training program on an age dummy, controlling for industry dummies, occupation dummies, year dummies, and gender within the sample of workers with vocational training.[40] Table 9, columns 1 and 2, shows the results. I find that workers below the age of 30 are indeed between 4.4 and 5.6 percentage points (pp) more likely to mainly work with computers and

[40]Evaluating whether trainees work more often with new technologies than workers who have completed their training program is not possible due to sample size issues.

Table 8: Use of vocational training according to firm survey

	Applies	Does not apply
Ensures supply of new skills and knowledge	51%	16%
Improves adaptability to technical change	46%	19%
Enhances innovative capabilities	51%	18%

Notes: Based on the BIBB-Cost-Benefit Survey 2000. Firms in East Germany only. On a scale from 1 ("Does not apply at all") to 5 ("Fully applies"). Applies: categories 4+5. Does not apply: Categories 1+2. Using representative survey weights. N=553.

computer-controlled machines than workers aged 30 and above, compared to a mean usage of 34.9%.[41] For more information, see Appendix F.

Table 9: Outcome: Use of computer-controlled machines (0/100)

	Main results		External validity across education groups			
			Low-educ. with VT	Highly educ. with VT	No education	Tertiary educated
	(1)	(2)	(3)	(4)	(5)	(6)
Reference category: 18-29 years						
30+	-5.60***	-4.40***	-5.00***	-3.10**	-4.40***	-2.18
	(0.79)	(0.69)	(0.98)	(1.51)	(0.69)	(1.53)
Controls		X	X	X	X	X
Mean dep. variables	34.90	34.90	39.91	29.95	34.90	24.35
N	45,488	45,488	28,769	8,540	45,488	11,281

Notes: Based on the BIBB-BAuA Qualification and Career Survey. 1999, 2006 and 2012 waves. All regressions control for dummies for the respective survey wave. Controls include gender, occupations (353), industries (17). Heteroscedasticity-robust standard errors. Columns 1 and 2: Among workers who have completed their training program. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Hence, the complementarity between young workers and new technologies is not only visible in the specific setting of the educational reform but also has more general implications. This analysis can be further expanded upon to inform the external validity of the established relation between young workers and firm technology adoption by looking at other worker groups. To do so, I reassess whether young workers more often work with new technologies than older workers also within other education groups. Columns 3 and 4 show that the result is found both for low-educated and highly educated workers who have completed their training program, even though the magnitude and precision is reduced for the latter (smaller) group. The pattern is also found for workers without formal education (neither vocational training nor tertiary education, column 5), and for tertiary educated workers (column 6). Again, the latter estimate lacks statistical precision, likely related to the much smaller sample size. I conclude that the complementarity between young labor market entrants and new technology, while potentially enhanced in the vocational training system, is something that likely also holds in other educational settings.

7.3 Alternative channels

There are at least three alternative explanations for the complementarity between young labor market entrants and technology adoption other than their low opportunity costs and the large productivity gains associated with learning new skills. First, according to standard human

[41] In industries which heavily employ highly educated trainees, numbers are similar: 2.2pp in the retail sector, 6.7pp in business-related services, and 4.2pp in private and public services.

capital theory, human capital investments in young workers yield longer-term benefits in expectation (the "horizon" channel in Cavounidis & Lang, 2020). Second, young workers might generally possess more up-to-date tech skills. Third, incumbent workers may be less willing to reskill.

While all of these channels may play a role, they are unlikely to fully cause the observed investment decline because they don't mark a discontinuous change between training cohorts. Put differently, they are unable to explain why marginally older trainees from the previous training cohort cannot act as substitutes for entrants when it comes to technology adoption. The only dimension on which new labor market entrants are considerably different to second-year trainees is in their opportunity costs and expected payoffs to acquiring new skills, as noted in Cavounidis & Lang (2020). Indeed, the Cost-Benefit Surveys of Vocational Training show that firm revenues from skilled labor activities of second-year trainees (third-year trainees) are 134% (254%) higher than for first-year trainees (Schönfeld et al., 2016, Table 18).

8 Discussion

In this paper, I provide empirical evidence that a temporary drop in the supply of vocational trainees causally reduces firm investments, linked to a decrease in technology adoption. This finding suggests that young labor market entrants are complements, rather than substitutes, to firm technology adoption. This complementary relationship can be explained by entrants' low opportunity costs of acquiring new skills and/or high expected pay-offs attached to it. Consequently, when young labor market entrants are scarce, firms face higher capital adjustment costs of worker training, reducing the adoption of technologies requiring new skills.

While it has been known that labor supply affects technology adoption, the crucial role of young labor market entrants and the mechanism via capital adjustment costs of training new skills are novel and informative from a number of perspectives: First, it highlights that the availability of young workers is a key factor determining firm technology adoption. While a reduction in the supply of young labor market entrants may not always cause a decrease in technology investments, it will always entail an increase in the costs of technology adoption. Second, it stresses the relevance of new skills demanded by new technologies. Third, it shows that retraining incumbents is costly, especially compared to training young labor market entrants. The combination of points two and three lead to strong vintage effects: worker cohorts posses different, vintage-specific skills.

While these general implications and the proposed mechanism via the comparative advantage of young labor market entrants in learning new skills is likely to hold in a broad range of settings, external validity hinges on the type of technology, the education system, the functioning of the labor market, and the duration of the supply reduction: First, the more productive a technology, and the fewer new skills required, the more likely the new technology will be implemented despite a shortage of young labor market entrants. Second, while young workers have lower opportunity costs of training than incumbents in most conceivable settings, the size of the effect may be larger in the context of German vocational trainees than in other contexts because the German vocational training system enhances skill transfer due to nationally binding training curricula and accompanying courses in vocational schools. Third, a negative supply shock of young labor market entrants can be absorbed by the labor market in different ways. If, for example, wages adjust such that employment of young labor market entrants does not decrease, the effect on technology adoption will be different. Last, the effect of a temporary supply shock likely differs from the effect of a long-term reduction due to general equilibrium effects. These aspects can explain the seemingly opposing finding by Abeliansky & Prettner (2017); Acemoglu & Restrepo (2022) that population aging increases the adoption of automation technologies; a setting looking at technologies that substitutes human tasks and not focusing on the availability of young labor market entrants.

From a policy perspective, the findings stress the importance of attracting young labor market entrants or subsidizing the retraining of incumbent workers to foster technology adoption, in particular in times of demographic change. The results also have implications for the optimal design of the German vocational training system: While the current system seems to effectively foster the adoption of new technologies, as indicated by Schultheiss & Backes-Gellner (2022), the finding that firms shy away from retraining incumbent workers who were trained a few years ago indicates that skills acquired through vocational training may be overly specific (compare Hanushek et al., 2017).

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., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. In Handbook of labor economics (Vol. 4, pp. 1043–1171). Elsevier.

Acemoglu, D., & Pischke, J.-S. (1998). Why do firms train? Theory and evidence. The Quarterly Journal of Economics, 113(1), 79–119.

Acemoglu, D., & Pischke, J.-S. (1999a). Beyond becker: Training in imperfect labour markets. The Economic Journal, 109(453), 112–142.

Acemoglu, D., & Pischke, J.-S. (1999b). The structure of wages and investment in general training. Journal of Political Economy, 107(3), 539–572.

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. (2024). Fast and slow technological transitions. Journal of Political Economy Macroeconomics, 2(2), 000–000.

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.

Arntz, M., Lipowski, C., Neidhöfer, G., & Zierahn, U. (2025). Computers as stepping stones? Technological change and equality of labor market opportunities. Journal of Labor Economics, forthcoming.

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., & Seegmiller, B. (2024). New frontiers: The origins and content of new work, 1940–2018. The Quarterly Journal of Economics, qjae008.

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.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. Journal of Political Economy, 70(5, Part 2), 9–49.

Becker, G. S. (1964). Human capital: A theoretical and empirical analysis, with special reference to education. Columbia University Press, New York.

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).

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.

Caicedo, S., Espinosa, M., & Seibold, A. (2022). Unwilling to train? Firm responses to the colombian apprenticeship regulation. Econometrica, 90(2), 507–550.

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

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

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

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

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

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

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

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. IAB-Discussion Paper, 28/2020.

Dustmann, C., & Schönberg, U. (2009). Training and union wages. The Review of Economics and Statistics, 91(2), 363–376.

Dustmann, C., & Schönberg, U. (2012). What makes firm-based vocational training schemes successful? The role of commitment. American Economic Journal: Applied Economics, 4(2), 36–61.

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, G.-O. (2022). 12411-0010: Bevölkerung: Bundesländer, Stichtag. (January 18, 2024; data license by-2-0; own calculation/own presentation.)

Federal Statistical Office, G.-O. (2023a). 13211-0007: Arbeitslose, Arbeitslosenquoten, gemeldete Arbeitsstellen: Bundesländer, Jahre. (January 18, 2024; data license by-2-0; own calculation/own presentation.)

Federal Statistical Office, G.-O. (2023b). 21711-0010: Ausgaben der öffentlichen Haushalte für Bildung: Bundesländer, Jahre, Körperschaftsgruppen, Aufgabenbereiche der öffentlichen Haushalte. (January 18, 2024; data license by-2-0; own calculation/own presentation.)

Federal Statistical Office, G.-O. (2023c). 71141-0006: Investitionsausgaben der öffentlichen Haushalte: Bundesländer, Jhre, Körperschaftsgruppen, Art der Investitionsausgaben [ohne Krankenhäuser und Hochschulkliniken mit kaufmännischem Rechnungswesen.]. (January 18, 2024; data license by-2-0; own calculation/own presentation.)

Federal Statistical Office, G.-O. (2023d). 71321-0002: Schulden der öffentlichen Gesamthaushalte: Bundesländer, Stichtag (31.12.1992-31.12.2005), Körperschaftsgruppen, Schuldenarten. (January 18, 2024; data license by-2-0; own calculation/own presentation.)

Federal Statistical Office, G.-O. (2023e). 82111-0001: Vgr der Länder (Entstehungsrechnung) – Bruttoinlandsprodukt zu Marktpreisen (nominal): Bundesländer, Jahre; i. jew. Preisen (Mill. Euro). (January 18, 2024; data license by-2-0; own calculation/own presentation.)

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.)

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

Fitzenberger, B., Osikominu, A., & Völter, R. (2006). Imputation rules to improve the education variable in the iab employment subsample. Journal of Contextual Economics–Schmollers Jahrbuch(3), 405–436.

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

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

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

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

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.

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

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.

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.

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

Lerche, A. (2022). Investment tax credits and the response of firms. IAB-Discussion Paper, 28/2022. doi: 10.48720/IAB.DP.2228

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

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

Lipowski, C., Salomons, A., & Zierahn-Weilage, U. (2024). Expertise at work: New technologies, new skills, and worker impacts. ZEW-Centre for European Economic Research Discussion Paper, 24-044.

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

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.

Moen, E. R., & Rosén, Å. (2004). Does poaching distort training? The Review of Economic Studies, 71(4), 1143–1162.

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

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

Müller, S. (2008). Capital stock approximation using firm level panel data: A modified perpetual inventory approach. Jahrbücher für Nationalökonomie und Statistik, 228(4), 357–371.

Müller, S. (2017). Capital stock approximation with the perpetual inventory method: An update. FDZ-Methodenreport, 5, 2017.

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

Rohrbach-Schmidt, D., & Tiemann, M. (2013). Changes in workplace tasks in Germany — evaluating skill and task measures. Journal for Labour Market Research, 46(3), 215–237.

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, 235(2), 2218–2244.

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. (2024). Are executives in short supply? Evidence from death events. Review of Economic Studies, 91(1), 519–559.

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.

Siegloch, S., Wehrhöfer, N., & Etzel, T. (2024). Spillover, efficiency and equity effects of regional firm subsidies. American Economic Jounral: Economic Policy, forthcoming.

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

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

Thomsen, U., Ludsteck, J., Schmucker, A., et al. (2018). Skilled or unskilled-improving the information on qualification for employee data in the iab employee biography. FDZ-Methodenreport, 9(2018), 22.

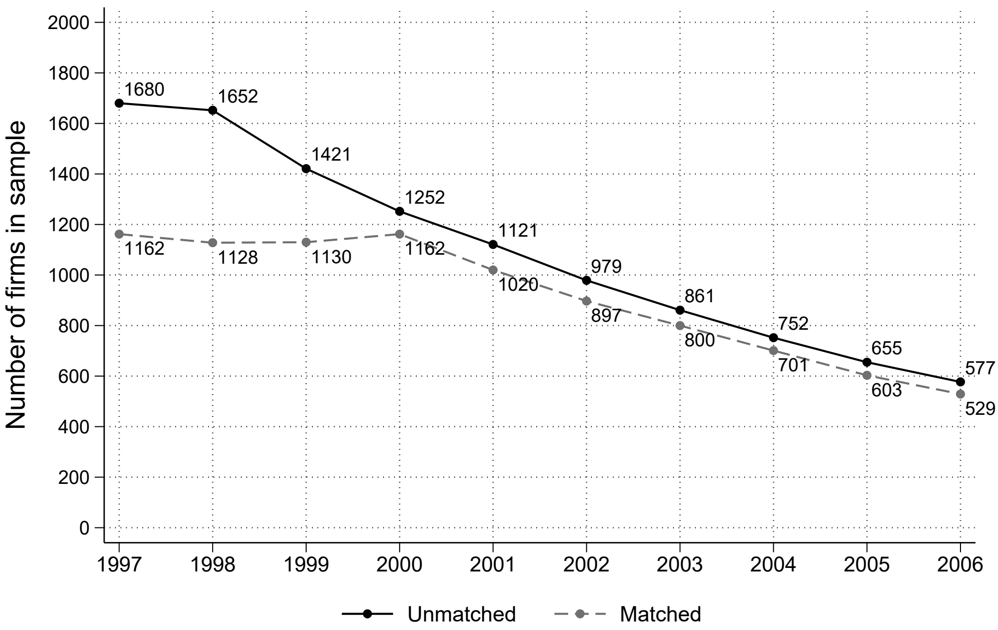
Walden, G., Beicht, U., & Herget, H. (2009). BIBB-Cost-Benefit-Survey 2000. doi: 10.7803/370.00.1.2.10

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

Figure A1: Panel attrition



Notes: Number of firms observed per year.

Table A1: Survey items used for the investment and technology indicators

Variable	Survey Question	Manipulation	Frequency
Inv. per worker	What was the approximate sum of all investments in t?	Divided by number of workers in 1997 from the administrative records. Trimming uppest percentile of the investment distribution and the investment per worker distribution	Yearly
Inv. type (0/1)	Did your establishment invest in one or more of the following areas in the last business year of t? EDP, information and communication technology? Production facilities, plant and equipment, furniture and fixture? Means of transport, transportation systems? Real estate and buildings?		Yearly
Technical status of machinery	How do you assess the overall technical status of the plant and machinery, furniture 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 organizational 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

Notes: t: Year of the survey. EDP: Electronic data processing.

B Potential reform confounders

Governing party – Social democrats. The education reform was an initiative of the Social Democratic Party, which entered the government in both treated states in 1994. In Mecklenburg-Western Pomerania, the Social Democrats were the junior governing party of a government led by the Christian Democratic Union. In Saxony-Anhalt, the were the senior governing party but shared power with the Greens. The Social Democrats also entered the government in one of the control states, Thuringia, in 1994 together with the Christian Democratic Union. Nonetheless, to exclude that the governance of the Social Democrats or other policy or socio-economic changes confound the effect of the education reform, I compare several state metrics including population size, education expenditure, unemployment rate, GDP, public debt and public investments between treated and control states before and after the reform, as well as between states governed by the Social Democrats and those not governed by the Social Democrats, see Table B1. Controlling for state and year fixed effects, I find no difference in any of these metrics between treated and control states post-reform compared to pre-reform that is statistically significantly different from zero, see Panel A. Turning to factors correlated with the governance of the Social Democrats, see Panel B, there is a significantly positive association between government of the Social Democrats and three indicators: education expenditure in % of the total state budget, unemployment rate, and log public investments. Higher education expenditure and public investments should, however, rather *in*crease instead of *de*crease firms technology investments. Regarding the unemployment rate, the relation to firm technology adoption is ambiguous. I conclude that major trends at the state level, potentially governed by the party composition of the government, are unlikely to cause the investment drop.

Investment subsidy programs. Two investment subsidy programs were in place in East Germany at the time that might have confounded the effects of the trainee supply shock. Below, I discuss each of them and how they may correlate with the trainee supply shock.

First, an investment tax credit policy was introduced in 1991 that aimed at supporting firms in former East Germany. In 1999, a policy change increased the tax credit rate for smaller manufacturing firms (with up to 250 employees) from 10% to 20%, and for larger firms from 5% to 10%, thereby reducing capital costs more significantly for smaller firms. Lerche (2022) exploits this reform and finds important increases in investments and employment in smaller compared to large firms in response to the reform. This reform is, however, unlikely, to confound the effect of the trainee supply shock given that my identification strategy relies on comparing firms across federal states but the tax credit reform had no regional variation.

Second, Germany's main regional policy, GRW, aimed at revitalizing underdeveloped regions, particularly East Germany, through investment subsidies for (mainly) manufacturing plants. The maximum subsidy rate varied based on counties' economic performance indicators, and was frequently reformed between 1997 and 2014. Siegloch et al. (2024) exploit these

Table B1: Correlation of state metrics with reform and Social Democratic Party

	(1) Log(Population)	(2) Log(Educ. expenditure)	(3) % education expenditure	(4) Unemployment rate	(5) Log (GDP)	(6) Log(Public Debt)	(7) Log(Public Investments)
				A. Education reform in 2001			
Treat × Post	-0.03	-0.01	1.37	-0.25	-0.00	0.15	0.15
	(0.11)	(0.19)	(2.46)	(0.84)	(0.17)	(0.32)	(0.15)
				B. Social democratic party in government			
Social Democrats	-0.01	0.05	0.94*	0.49*	-0.00	0.05	0.19**
	(0.01)	(0.04)	(0.50)	(0.28)	(0.02)	(0.04)	(0.07)
Mean dep. variable	14.83	21.52	27.68	18.08	10.74	9.21	6.09
N	84	66	66	84	84	84	84

Notes: Panel A: Treated: Mecklenburg-Western Pomerania and Saxony-Anhalt. Post: 2001 onward. *Panel B:* Social democrats among governing parties (1/0). Controlling for state and year fixed effects. Observations at the state-year level for East German states for 1992 until 2005, except for education expenditure (column 2 and 3) that is only observed from 1995 onward. Education expenditure: Total public expenditure on education. Share education expenditure: Public expenditure on education as a percentage of the total budget. Unemployment rate: Unemployment rate in % of the dependent civilian labor force. Debt: Debt of the overall public budget. Sources: (1) – Federal Statistical Office (2022) (2) & (3) – Federal Statistical Office (2023b) (4) – Federal Statistical Office (2023a) (5) – Federal Statistical Office (2023e) (6) – Federal Statistical Office (2023d) (7) – Federal Statistical Office (2023c) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

reforms to study the effect of tax credit on investment and employment, finding important effects. This program had a regional component: while all counties were assigned the same maximum subsidy threshold in 1990, 27 counties that were previously assigned as high-subsidy counties were assigned as low-subsidy county in 1997. In 2000, nine further counties were assigned from high to low, while one was assigned from low to high.

Studying the regional correlation between this potentially confounding reform and the 2001 education reform suggests that the 1997 changes are unlikely to have caused the investment decline in Mecklenburg-Western Pomerania and Saxony-Anhalt following the 2001 trainee shortage: in Brandenburg, half of all counties were downgraded; Saxony saw 38% of its counties affected, and Thuringia experienced a change in 35% of its counties. In contrast, the treated states, Mecklenburg-Western Pomerania and Saxony-Anhalt, were the least affected, with only 25% and 21% of their counties downgraded, respectively. This should have led to increased investments in these states relative to the control states.

The 2000 reform primarily impacted Saxony, where 39% of counties were downgraded from high-subsidy status to low-subsidy status. Brandenburg and Western-Pomerania were not affected at all. Thuringia saw a downgrade in 9% of its counties, and Saxony-Anhalt had a net change of 7% with 14% of the counties changing from high status to low status, and 7% from low status to high status. Therefore, the decline in investments in Mecklenburg-Western Pomerania cannot be explained by these reforms, and it is very unlikely that the changes caused the decrease in Saxony-Anhalt.

C Additional results

Table C1: Correlation of the error term within states and within firms

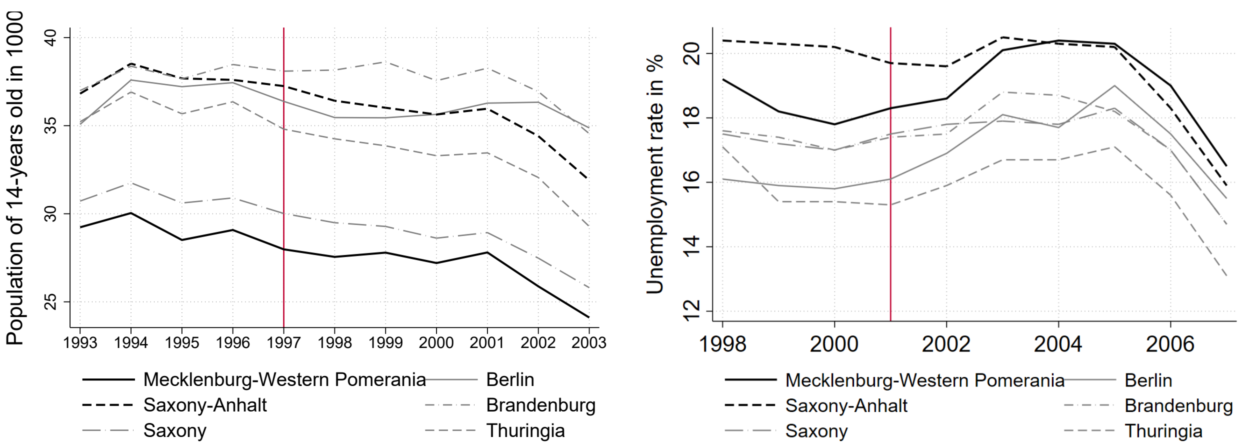
	Overall SD	SD within states	SD within firms
# highly educated trainees	5.10	4.23	2.40
Inv. per worker in €1,000	18.42	18.16	13.74
Adjusted log investments	3.01	3.00	2.41

Notes: Standard deviations (SD) of the error term resulting from a regression following equation (1) with the outcome variable shown in the first column. Adjusted log investments: $log(\text{inv})\,\forall\,\text{inv} > 0; log(\text{inv}) := -0.01\,\forall\,\text{inv} = 0$, reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin.

Figure C1: Demographic and economic trends across federal states

A. Number of 14-years-old by state

B. Unemployment rate by state



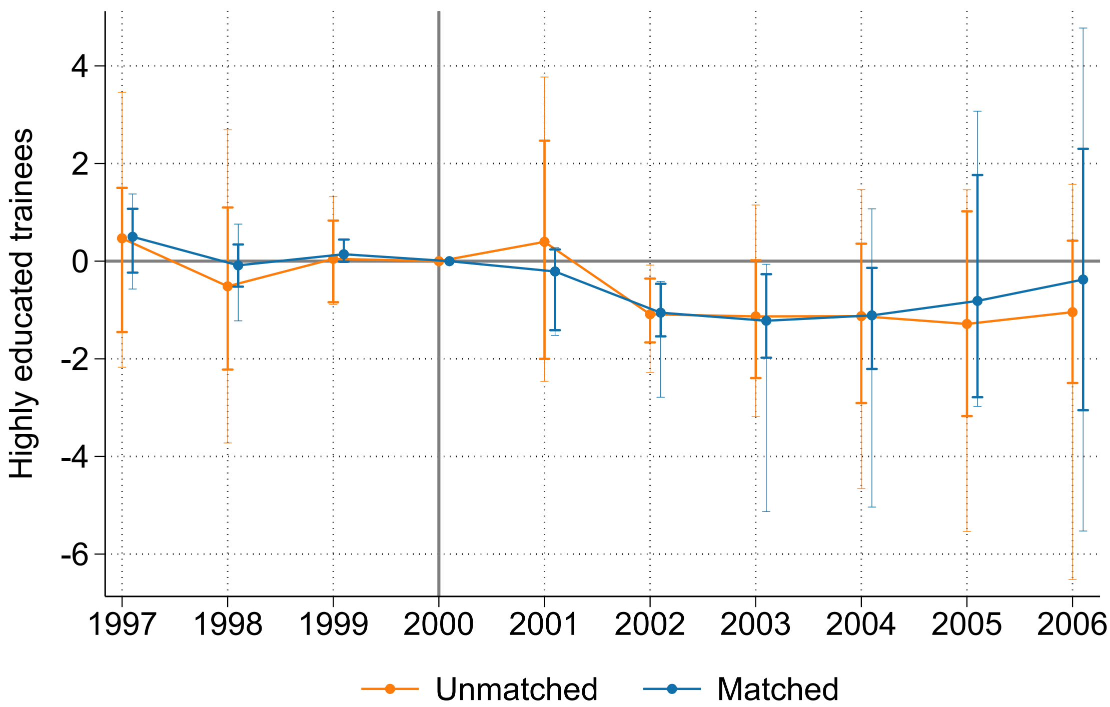
Notes: Panel A: Source: Federal Statistical Office, Genesis-Online (2022b). The number for Saxony is divided by two for better visibility. *Panel B:* Source: Federal Statistical Office, Genesis-Online (2022b).

Table C2: Summary statistics of outcome variables

	Mean	SD	Within-firm SD	Min	Max	N
# highly educated trainees	5.45	14.83	4.59	0.00	461.00	4678
Adjusted log investments	12.37	5.09	2.97	-0.01	17.24	4678
Inv. per worker in €1,000	16.76	31.26	16.07	0.00	663.27	4678
Technical status (1–5)	3.90	0.72	0.47	1.00	5.00	4656
Organizational change (0–4)	1.02	1.16	0.75	0.00	4.00	2245
Log(capital stock)	10.18	2.38	0.30	1.58	13.49	4393
Log(employment)	5.06	1.32	0.34	0.00	9.40	4678

Notes: SD: Standard deviation. N: number of observations. Among training firms, over the time period 1997–2006. Adjusted log investments: $log(\text{inv})\,\forall\,\text{inv} > 0; log(\text{inv}) := -0.01\,\forall\,\text{inv} = 0$, reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin.

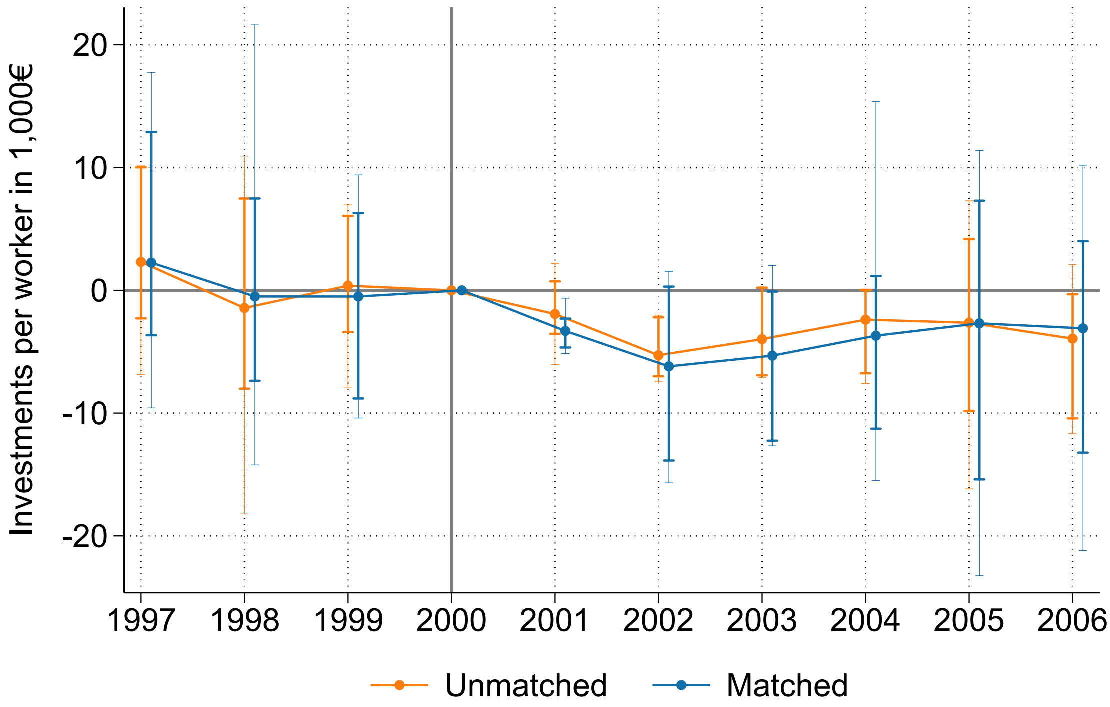
Figure C2: Trainee employment effects – Cluster wild t-bootstrap confidence intervals



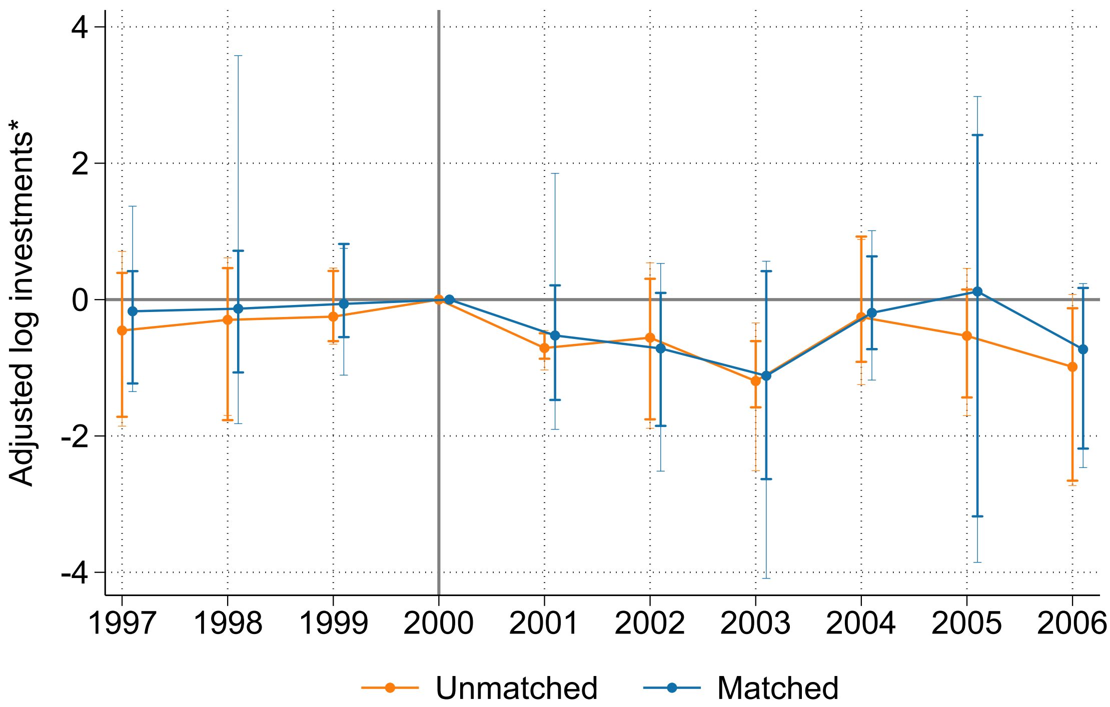
Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (1). Confidence intervals based on cluster wild t-bootstraps following Cameron et al. (2008). Among training firms only. For the main figure, see Figure 3. N=3,322 for the unmatched sample and N=3,182 for the matched sample.

Figure C3: Effect on investments – Cluster wild t-bootstrap confidence intervals

A. Investments per worker in €1,000

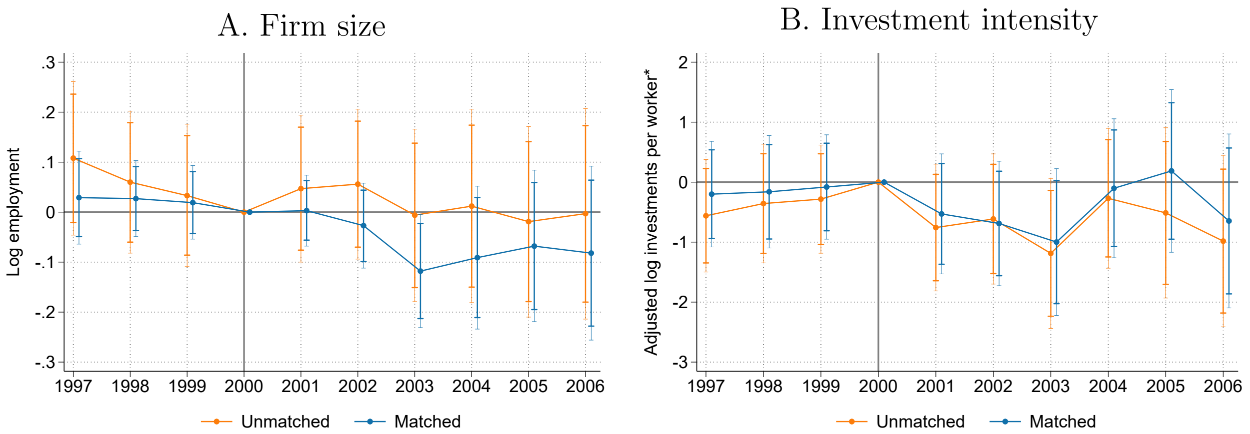


B. Adjusted log investments



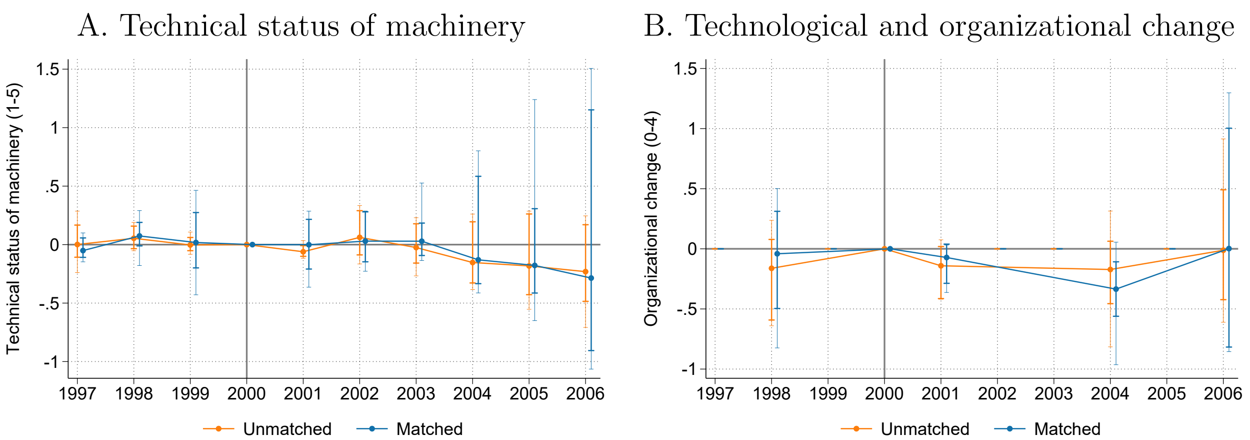
Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (1). Confidence intervals based on cluster wild t-bootstraps following Cameron et al. (2008). Among training firms only. Outcome Panel A: investments in €1,000 divided by total employment in 1997. Outcome Panel B: Adjusted log investments: $log(\text{inv})\,\forall\,\text{inv} > 0; log(\text{inv}) := -0.01\,\forall\,\text{inv} = 0$, reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin. For the main figures, see Figure 4. N=3,322 for the unmatched sample. N=3,182 for the matched sample.

Figure C4: Effect on firm size versus investment intensity



Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (1). Standard errors clustered at the firm level. Among training firms only.

Figure C5: Effect on technology adoption – Cluster wild t-bootstrap confidence intervals



Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (1). Confidence intervals based on cluster wild t-bootstraps following Cameron et al. (2008). Among training firms only. For the main figures, see Figure 7.

Table C3: DiD Results - Wage and worker substitution effects (Full table)

	Log wage highly educ. trainees	# low-educ. trainees	# highly educ. commuting trainees	Log highly educ. VT employment	Log wages educ. VT employment	Trainee retention rate	Internal retraining
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				A. All training firms (Unmatched)			
Treat × Roll-out	-0.00	-0.72	0.83	0.06	0.02	-0.03	0.03
	(0.02)	(1.11)	(2.73)	(0.05)	(0.02)	(0.05)	(0.21)
Treat × Post	-0.00	-0.90	2.46	0.01	-0.02	-0.06	-0.27*
	(0.03)	(1.77)	(3.18)	(0.08)	(0.02)	(0.04)	(0.16)
Treat × Phase-out	0.04	0.62	1.77	-0.08	0.01	-0.03	-2.05
	(0.05)	(1.31)	(3.79)	(0.12)	(0.03)	(0.05)	(1.91)
Mean dep. variable	3.02	10.06	2.83	2.18	4.29	0.65	0.47
N	2252	3322	1429	3083	3082	3150	1618
				B. Matched training firms			
Treat × Roll-out	0.01	-2.13	-0.38	0.07	0.02	-0.03	-0.16**
	(0.02)	(1.31)	(2.32)	(0.05)	(0.02)	(0.05)	(0.08)
Treat × Post	0.01	-1.65	2.04	0.03	-0.02	-0.09**	-0.09
	(0.03)	(1.76)	(2.87)	(0.09)	(0.03)	(0.04)	(0.07)
Treat × Phase-out	0.05	-1.56	0.16	0.01	0.02	-0.04	-0.13
	(0.05)	(1.31)	(4.13)	(0.14)	(0.03)	(0.06)	(0.11)
Mean dep. variable	3.03	9.72	2.93	2.25	4.31	0.64	0.50
N	2198	3182	1564	3032	3031	3035	1586

Notes: Difference-in-difference coefficients based on equation (2). Pre: 1997–2000. Roll-out: 2001. Post: 2002–2004. Fade-out: 2005–2006. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Among training firms only. Mean dep. variable: Average outcome of treated firms in 2000. Column 3: a commuter is defined as a person living and working in two distinct federal states. This variable is available from 1999 onward. Column 6: 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. Column 7: Internal retraining is the sum of retraining incidences at the firm-year level. VT: completed vocational training. For the main table, see Table 4.

Table C4: DiD Results – Investment effects (Full table)

	Log(K) (1)	Any inv. (0/1) (2)	Log(Inv.) (3)	Large inv. (1/0) (4)
		A. All training firms (Unmatched)		
Treat × Roll-Out	-0.04	-0.02	-0.06	-0.11**
	(0.04)	(0.04)	(0.16)	(0.05)
Treat × Post	-0.07***	-0.02	-0.16	-0.11**
	(0.05)	(0.04)	(0.15)	(0.05)
Treat × Phase-Out	-0.08	-0.02	-0.06	-0.07
	(0.07)	(0.03)	(0.16)	(0.05)
Mean dep. variable	10.18	0.90	13.98	0.33
N	3155	3308	2843	2843
		B. Matched training firms		
Treat × Roll-Out	-0.04	-0.01	-0.14	-0.14**
	(0.03)	(0.05)	(0.18)	(0.05)
Treat × Post	-0.10*	-0.03	-0.24	-0.16***
	(0.06)	(0.04)	(0.16)	(0.05)
Treat × Phase-Out-Out	-0.10	-0.01	-0.04	-0.06
	(0.08)	(0.04)	(0.19)	(0.06)
Mean dep. variable	10.04	0.89	13.82	0.30
N	3064	3176	2809	2809

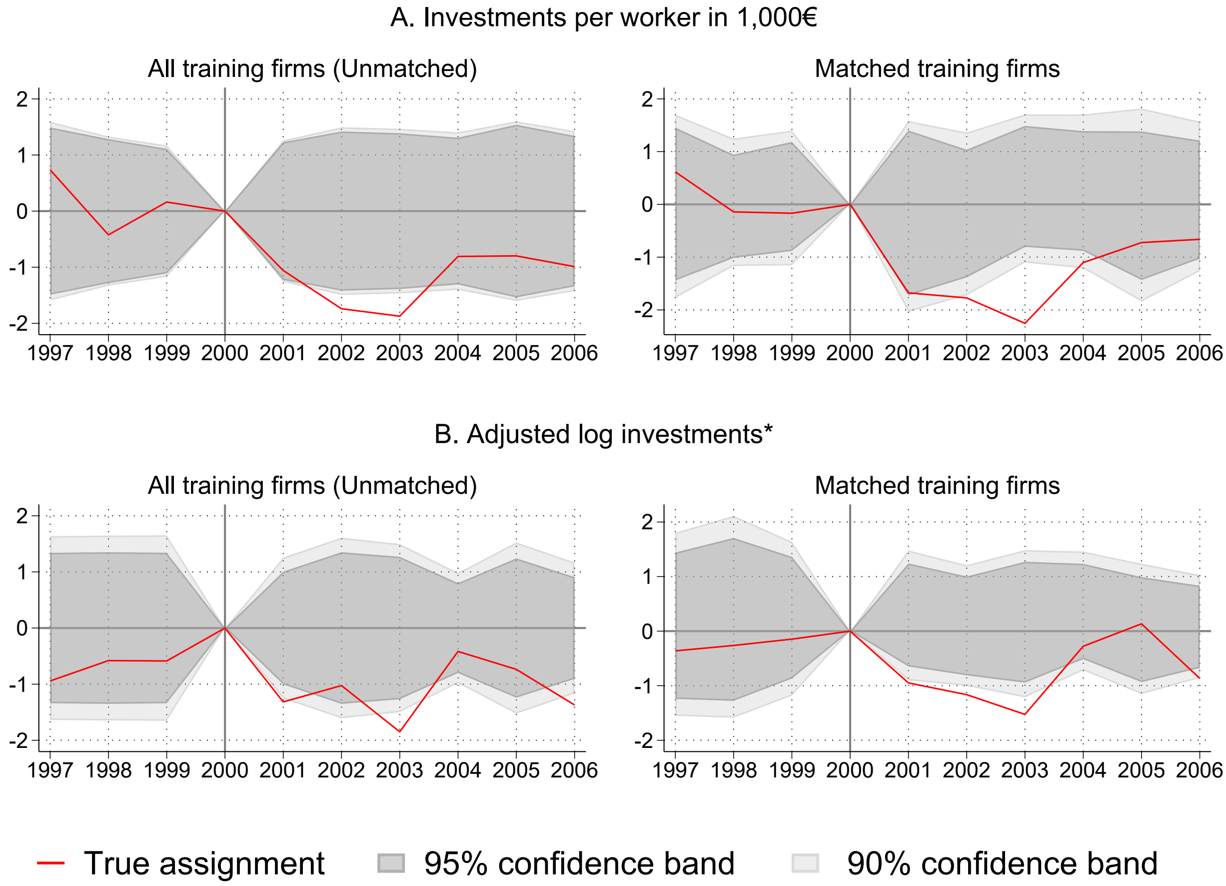
Notes: Pre: 1997–2000. Roll-out: 2001. Post: 2002–2004. Fade-out: 2005–2006. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mean dep. variable: Average outcome of treated firms in 2000. For the main table, see Table 5.

Table C5: DiD Results – Investment effect heterogeneity – Matched sample

	Firm size		Industry		
	Small	Big	Business services	Manufacturing	Other
	(1)	(2)	(3)	(4)	(5)
			A. Investments per worker		
Treated × Post	-8.17	-3.60*	-7.90	-5.72**	-0.43
	(5.82)	(1.86)	(5.02)	(2.71)	(2.89)
Mean dep. variable	17.34	16.72	22.87	13.15	6.34
			B. Adjusted log investments		
Treated × Post	-0.91	-0.62	-0.56	-1.86	-0.99
	(1.00)	(0.49)	(0.64)	(1.19)	(1.53)
Mean dep. variable	10.63	14.47	13.19	12.88	10.22
N	1572	1610	1812	686	684

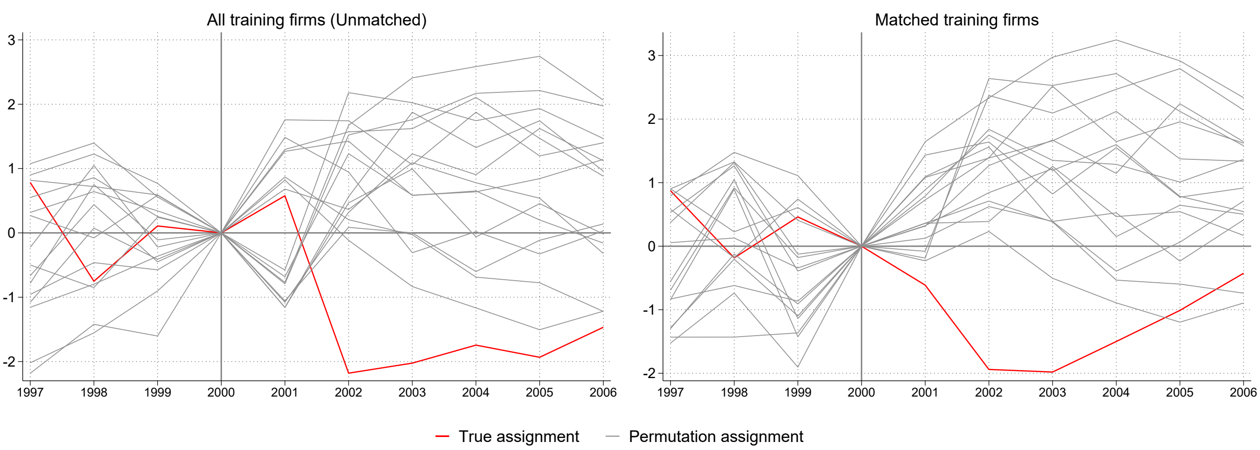
Notes: Difference-in-difference coefficients based on equation (2). Pre: 1997–2000. Roll-out: 2001. Post: 2002–2003. Fade-out: 2004–2006. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Among training firms only. Mean dep. variable: Average outcome of treated firms in 2000. Small firms: median or below median overall employment in 1997. Large firms: above median overall employment in 1997. Business services: Firms operating in business services and public administration.

Figure C6: Permutation test in West Germany – Effect on investments



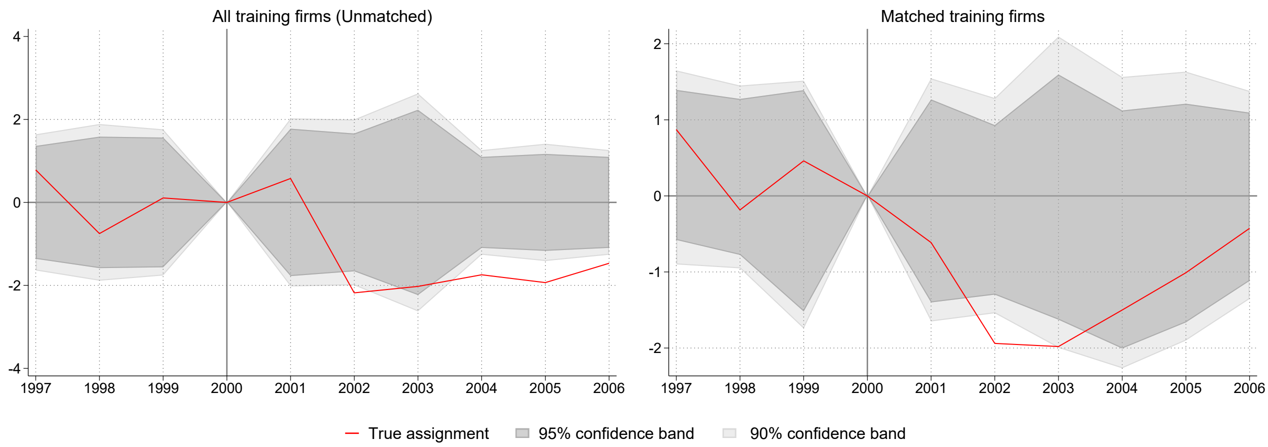
Notes: T-statistics of the difference-in-differences event study coefficients based on equation (1) using the actual treatment assignment (red line) and all possible permutation assignments across West German states (gray areas).

Figure C7: Permutation test – Effect on trainee employment



Notes: T-statistics of the difference-in-differences event study coefficients based on equation (1) using the actual treatment assignment (red line) and all possible permutation assignments (gray lines).

Figure C8: Permutation test in West Germany – Effect on trainee employment



Notes: T-statistics of the difference-in-differences event study coefficients based on equation (1) using the actual treatment assignment (red line) and all possible permutation assignments across West German states (gray areas).

D Instrumental variable regression

In this Appendix, I examine the treatment effect along the intensive treatment margin using a complementary identification strategy. This analysis serves three main purposes. First, it allows me to understand whether firms that are more affected by the trainee supply reduction indeed decrease investments more. Second, by only leveraging the exogenous part of the treatment intensity using an instrumental variable, it allows me to identify the treatment effect independent of the realized, and potentially endogenous, distribution of trainees across firms. Third, the analysis hereby identifies a different causal parameter: While the event study approach identifies the causal effect of facing a statewide reduction in trainee supply, this complementary analysis identifies the causal effect of one additional trainee.

I estimate a two-stage-least-squares (2SLS) model of firm investments Inv on firm employment of highly educated trainees N^{Trainee} controlling for firm fixed effects π_j and year fixed effects ψ_t, see equation (D1). I instrument trainee employment as given in equation (D2):

$$\text{Inv}_{jt} = N_{jt}^{\text{Trainee}} + \psi_t + \pi_j + \epsilon_{jt} \tag{D1}$$

$$\begin{aligned} N_{jt}^{\text{Trainee}} &= \sum_t \gamma_t (N_{j,1997/98}^{\text{Trainee}} \times \text{Treat}_j \times \text{Year}_t) \\ &+ \sum_t \zeta_t (N_{j,1997/98}^{\text{Trainee}} \times \text{Year}_t) + \psi_t + \pi_j + \epsilon_{jt} \end{aligned} \tag{D2}$$

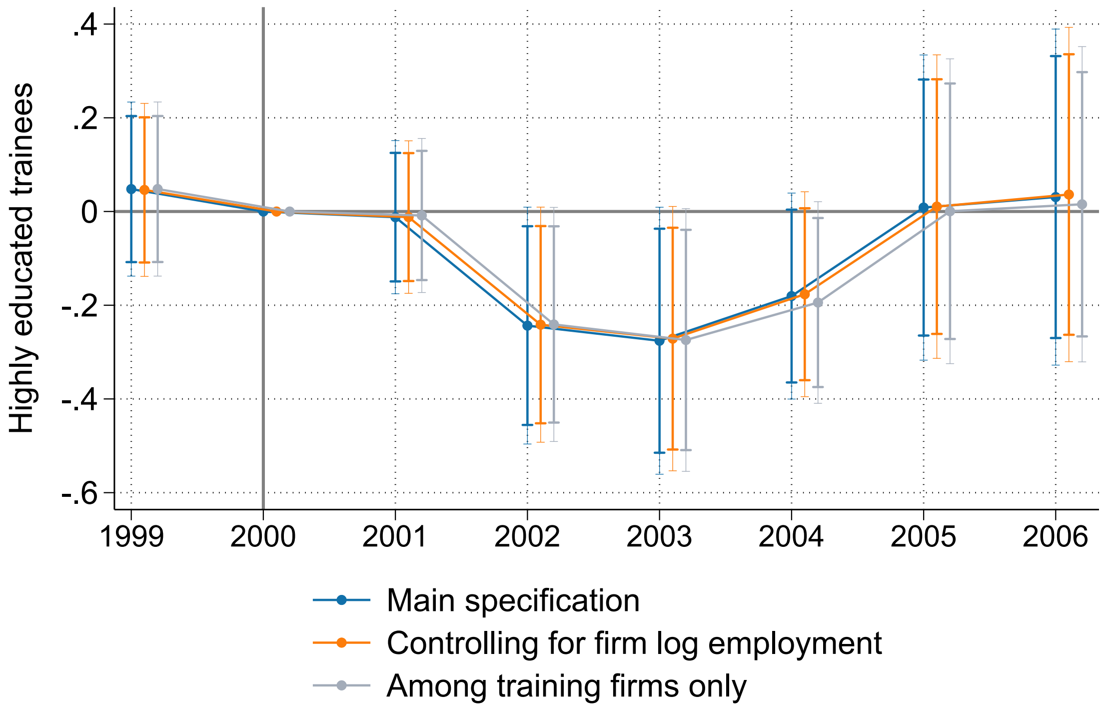
with j firms, and t calendar years. Treat takes the value one if a firm is located in a state undergoing the education reform and zero otherwise. I predict contemporaneous trainee employment by firms' initial mean employment of highly educated trainees in 1997/1998, i.e. firm exposure, $N_{1997/98}^{\text{Trainee}}$, corresponding to the *shares* in a shift-share instrument, times Treat $\times$ Year, corresponding to the reform-induced *shifts* in the supply of trainees across states and years. I control for time trends in firm trainee employment which are allowed to vary by firm exposure, $N_{j,1997/98}^{\text{Trainee}} \times \text{Year}_t$. Hence, the instrument exploits variation between two equally exposed firms located in a treated state and a control state across time. The exogeneity of the instrument stems from the random assignment of the trainee supply shock, i.e. the education reform, to states and years. Since employment of highly educated trainees in 1997/98 is expected to directly impact investments of the same year, which would violate the exclusion restriction, I run the regression for the years 1999 onward.[42] I estimate the effect within the sample of matched firms,[43] ensuring that treated and control firms are comparable in terms of sector, size, and employment of highly educated trainees.

Figure D1 shows the coefficients of interest of the first stage, γ_t. One additional highly educated trainee prior to the reform is associated with 0.24–0.28 fewer highly educated trainees in 2002 and 2003, and 0.18 fewer highly educated trainees in 2004. These estimates are smaller

[42] Results are robust to further restricting to the years 2000 onward.

[43] For consistency, I employ the same matching procedure as above, i.e. matching treated training firms to control training firms and treated non-training firms to control non-training firms.

Figure D1: IV results – First stage



Notes: Coefficients plus 90% and 95% confidence intervals of the term $(N^{\text{Trainee}}_{j,1997/98} \times \text{Treat}_j \times \text{Year}_t)$ in equation (D2). Outcome: Number of highly educated trainees. Standard errors clustered at the firm level.

than a third—the theoretical number if one out of three training cohorts was missing—but slightly larger than those implied by the event study design (-0.24 in 2002 based on the IV approach compared to -1.06/6.17=-0.17). Consistent with the timing of the shock, the coefficients in 2005 and 2006 are zero. The first stage thus confirms that initial trainee exposure is a relevant instrument. With F-statistics of approximately 15 to 16, see Table D1, Panel A, the instrument is relevant.

Table D1, Panel A, also shows the results of the second stage for different specifications of the investment outcome. While only borderline significant, the positive coefficients imply that a reduction in trainee employment decreases firm investments. In particular, one fewer trainee reduces investments by €930 per worker (column 1), and the capital stock by approximately 2% (column 4). The investment decline is driven by the intensive investment margin (column 3), in line with the results based on the difference-in-differences event study. Note that the estimate is identified for reductions in trainee employment of around 1.1 (as we know from the event study) but is unlikely to be linearly scalable for substantially larger drops.

To ensure that the relationship between trainees and investments is not (exclusively) driven by the role trainees play in firm employment growth, I control for time-variant log employment in a robustness check, see Panel B. Convincingly, the results remain very similar.

When restricting the sample to training firms only (Panel C), the estimates look comparably, but are mostly not statistically significant. I conclude that the distinction between employing

Table D1: IV results – Second stage

	Inv. per worker (1)	Adj. log inv.* (2)	Log inv. (3)	Log(K) (4)
	A. Main specification			
N^{Trainee}	0.93*	-0.09	0.04**	0.02**
	(0.53)	(0.06)	(0.02)	(0.01)
F-Stat	15.26	15.26	16.40	16.58
	B. Controlling for firm log employment			
N^{Trainee}	0.92*	-0.09	0.04**	0.02**
	(0.54)	(0.06)	(0.02)	(0.01)
F-Stat	15.41	15.41	16.71	16.78
N	7,037	7,037	5,207	6,737
	C. Among training firms only			
N^{Trainee}	0.61	0.02	0.04*	0.01**
	(0.47)	(0.05)	(0.02)	(0.01)
F-Stat	13.90	13.90	13.43	15.52
N	1,579	1,579	1,349	1,529

Notes: F-Stat gives the robust Kleibergen-Paap Wald rk F statistic. Based on the matched sample. Inv. per worker: investments in €1,000 divided by total employment in 1997. Adjusted log investments: $log(\text{inv})\,\forall\,\text{inv} > 0; log(\text{inv}) := -0.01\,\forall\,\text{inv} = 0$, reflecting that a change at the extensive margin is valued as much as a 1% change at the intensive margin. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

or not employing trainees is more relevant for firm investments than the number of trainees conditional on having at least one trainee.

In summary, while estimates turn partly imprecise and the F-statistics are not always as large as desired, the overall picture based on this complementary identification strategy confirms the negative impact of reduced trainee supply on firm investments and demonstrates that firms more affected by the negative trainee supply shock reduce investments more. The effects are smaller than the ones implied by the ratio between missing trainees and missing investments in the event study regression. This discrepancy might be due to spill-over effects within treated states, i.e. firms decreasing investments beyond the first-order decrease related to foregone trainee employment. It might also hint at firm selection into trainee employment: If firms that would have invested in the absence of the supply shock employ fewer trainees than firms who would not have invested anyway, the parameter identified in the event study approach is inflated, while the parameter identified in the IV approach is unaffected. A third reason may be trainee quality: firms employing trainees during the shortage also reduce investments because the trainees they employ are of worse quality.

E Economic framework

In this Appendix, I provide a formal exposition of the stylized economic framework laid down in Section 7.1. The model is implicitly grounded in the tasks framework à la Acemoglu & Autor (2011)—where new technologies substitute or complement labor in existing tasks, and introduce new tasks performed by human labor—but introduces technology vintages requiring vintage-specific skills, and capital adjustment costs of worker training in these new skills.

Setting. Suppose firms operate and employees work in overlapping generations. In each period t, each firm j produces one final good Y using labor L and vintages of production technologies τ with fixed marginal productivities A_τ under the following production function:

$$Y_{jt} = \sum_{\tau=0}^{\mathcal{T}} y_{j\tau} = \sum_{\tau=0}^{\mathcal{T}} A_\tau L_{jt\tau} \tag{E1}$$

For simplicity, I abstract from tasks performed by capital and I assume that the tasks $y_{j\tau}$ are perfect substitutes.[44] Each technology vintage requires specific skills such that only those workers who have been trained for the specific technology, L_τ, can use the new technology τ which is equivalent to performing the new task introduced by technology τ. The price for the final product is fixed to one for simplicity.

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

Firms decide whether to adopt the new technology at the start of the period in order to maximize profits. To adopt the new technology, firms may (re-)train a fraction Ψ_{τ_0} of workers of each initial productivity type τ_0. Training uniformly takes one period across technologies and workers.Since workers within a cohort are homogeneous, firms always either retrain all or no worker of each entry cohort, $\Psi_{\tau_0} = \{0; 1\}$, such that worker cohorts and worker skill types coincide. The wage rate w_τ is in proportion to, but below workers' 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.[45]

Costs of technology adoption consist of capital adjustment costs C that are equal to the costs of worker training. Training costs are born by the firm and are equal to the sum of foregone outputs of all workers undergoing training, with the training costs for each cohort

[44] While this assumption can be relaxed, it allows me to easily target *changes* in firm profits instead of total firm profits in the maximization problem below.

[45] The assumption that wages are not equal to marginal productivity is well backed up in the literature, in particular in the context of firm training (e.g. Konings & Vanormelingen, 2015).

given by:

$$C_{j\tau_0} = A_{\tau_0}\Psi_{\tau_0}L_{jt\tau_0} \tag{E2}$$

Firm maximization problem. For simplicity, assume firms maximize over two periods only, the training, $t = 1$, and the production period, $t = 2$, and workers do not change firms. Given the additive separability of tasks and the discrete nature of the adoption problem, firms maximize additional profits from technology adoption by deciding whether to adopt and train for each initial worker type $\tau_0 \in [0, ..., \mathcal{T} - 1]$ separately. Firms' additional profits for each initial worker type are equal to the net output surplus—defined as the output surplus minus the surplus in the wage bill—in period $t = 2$ minus capital adjustment costs of worker training. Hence, the firms maximization problem regarding worker type τ_0 can be written as

$$\max_{\Psi_{\tau_0}} \Delta Y_{j\tau_0} - \Delta W_{j\tau_0} - C_{j\tau_0} \tag{E3}$$

with

$$\Delta Y_{j\tau_0} - \Delta W_{j\tau_0} = (1-\theta)\Psi_{\tau_0}L_{jt\tau_0}(A_\tau - A_{\tau_0}) \tag{E4}$$

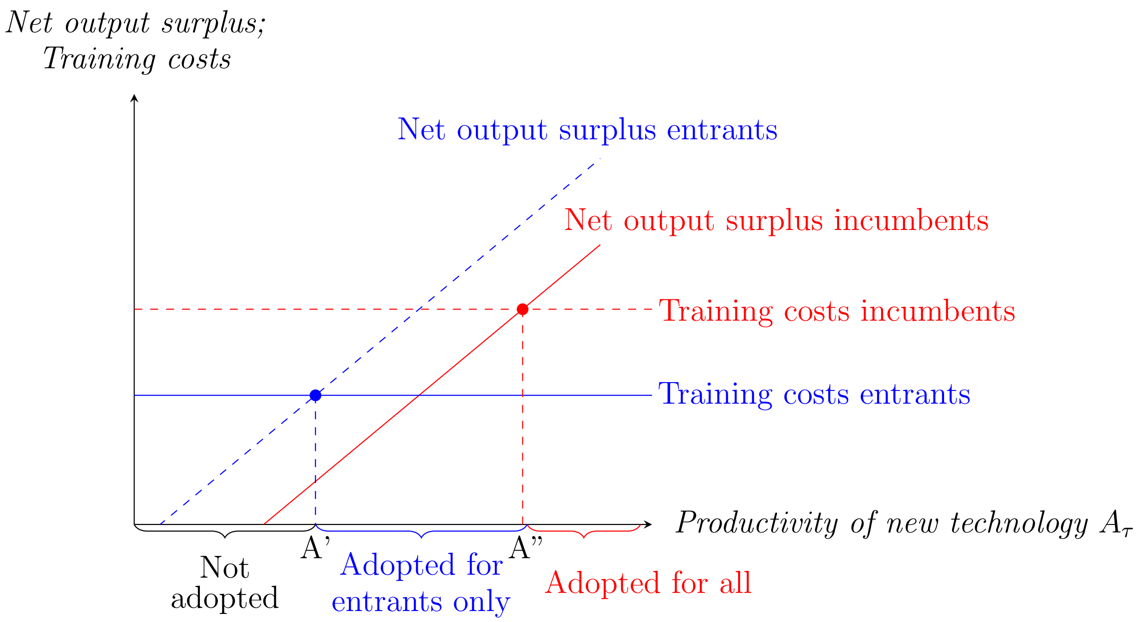
The profitability of training decreases in workers' initial productivities: The higher a worker's initial productivity, the lower the net output surplus and the higher the training costs. Combining equations (E1)–(E4), it follows that firms train an initial worker type L_{τ_0}, choosing $\Psi_{\tau 0} = 1$, as long as additional profits exceed additional costs, i.e. as long as the following condition between the productivity of the new technology, A_τ, and workers' initial productivity, A_{τ_0}, holds:

$$A_\tau \geq \left(1 + \frac{1}{1-\theta}\right) A_{\tau_0} \tag{E5}$$

and $\Psi_{\tau 0} = 0$ otherwise. Figure E1 visualizes this trade-off. New technologies below a productivity threshold A' are not adopted because training costs are too high, even for the least productive workers, i.e. the entrants of the current period. New technologies above A' but below a productivity threshold A'' are adopted by training labor market entrants only. New technologies above a threshold A'' are adopted by retraining incumbent workers as well.

Assume there is a missing entry cohort in $t = 1$. In this period, firms invest in the new technology τ if and only if the new technology is productive enough to make it profitable to retrain incumbents, i.e. if the condition in equation (E5) is satisfied for incumbents. If the new technology is of a productivity $A' \leq A_\tau < A''$, firms reduce their technology adoption compared to the case without a missing entry cohort. Highly productive technologies above A'' are always adopted, also in times of a shortage of young labor market entrants.

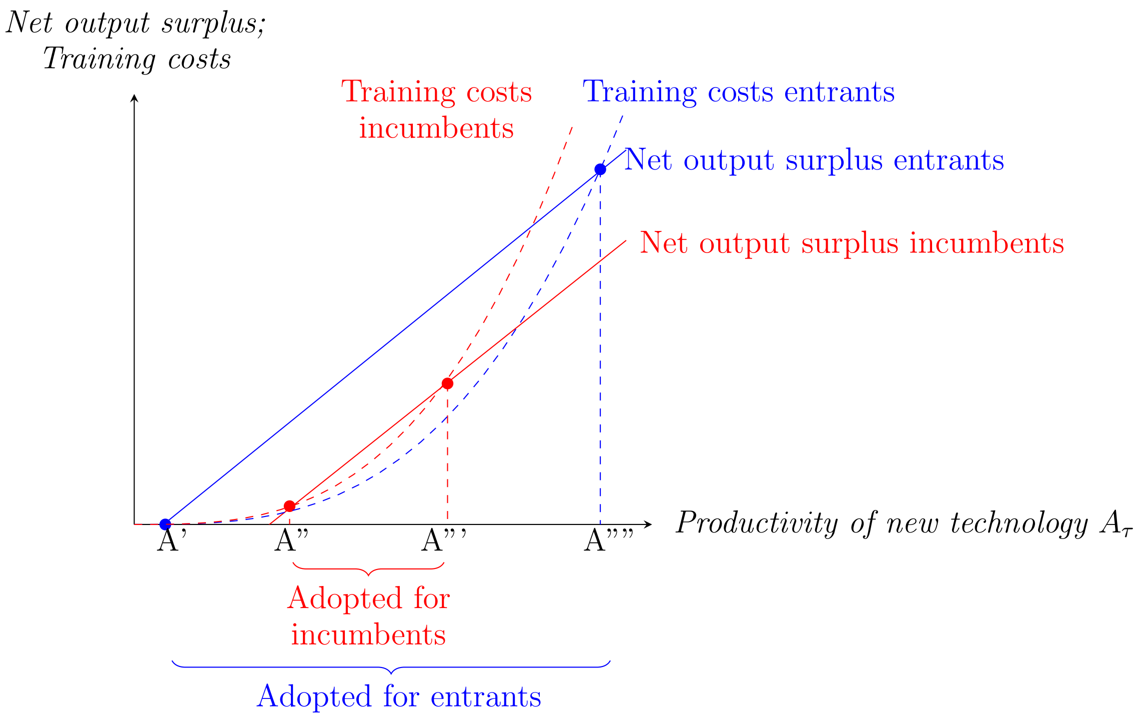
Figure E1: Additional profits versus additional costs of technology adoption



Notes: Profitability of training entrants versus incumbent workers. Net output surplus defined as the output surplus minus the surplus in the wage bill.

Extension – Increasing and convex capital adjustment costs. Until now, I have assumed training costs to be constant over all productivity levels. In standard capital adjustment costs models, adjustment costs are assumed to be increasing and convex in investment size. Let us now assume that training costs are increasing and convex in technology productivity, $C'(A_\tau) > 0, C''(A_\tau) > 0$. This implies that small investments can be easily incorporated in the structure of the firm without much training, while large investments create more pronounced disruptions requiring longer training. The trade-off between additional profits and additional costs of production is shown in Figure E2. The trade-off looks similar for small productivity levels of the new technology: New technologies below a productivity threshold A' are not adopted because training costs are too high; new technologies above A' but below A'' are adopted by training labor market entrants only; and new technologies above a threshold A'' are adopted by retraining incumbent workers as well. In addition, there are new, highly productive technologies above a certain productivity threshold A''' that require prohibitively long training to justify retraining incumbents, and even more productive technologies above a certain productivity threshold A'''' for which training costs exceed productivity gains for all workers. Consequently, a lack of entrants not only hinders the adoption of rather unproductive new technologies in the range between A' and A'', but also of very productive technologies with productivities in the range between A''' and A''''.

Figure E2: Additional profits versus additional costs with convex adjustment costs



Notes: Profitability of training entrants versus incumbent workers when capital adjustment costs of training are increasing and convex in technology productivity. Net output surplus defined as the output surplus minus the surplus in the wage bill.

F Supporting evidence based on additional datasets

BIBB Cost-Benefit Surveys. The Federal Institute for Vocational Education and Training (BIBB) surveys companies at intervals of several years on the benefits and costs of their vocational training, constituting a dataset called Cost-Benefit-Survey. A total of 2,518 companies took part in the 2000 survey. The data is representative of all German companies with training activities. The interviews are conducted with people who are primarily responsible for organizing and carrying out in-company training activities in these companies. For more information, see (Walden et al., 2009). For the analysis, I restrict the data to East German firms.

Among other things, firms are asked "To what extent do the following statements apply to your company's own training?" on a five point scale from "fully applies" to "does not apply at all", with three of 17 statements being "ensures the constant inflow of new knowledge into our company", "significantly improves our ability to adapt to technical and market changes" and "improves the innovative capacity of our company" (question 13 in the 2000 benefit questionnaire).

IAB/BIBB/BAuA Qualification and Career Surveys. This analysis is based on the 1999–2012 waves of the IAB/BIBB/BAuA Qualification and Career Survey (QCS). The QCS are repeated cross sectional surveys conducted by BIBB, IAB, and BAuA.[46] The survey covers around 30,000 employees. With the only difference not excluding trainees and East Germany, I closely follow Arntz et al. (2025) and I use the harmonization and data restriction procedure by Rohrbach-Schmidt & Tiemann (2013).

The surveys include questions regarding the main working tool used by each respondent. In the 1992 wave, these tools were categorized into (1) non-mechanical tools (e.g. handcart, pencil), (2) tools with some mechanization (e.g. telephone, hand drill machine), (3) tools with advanced mechanization (e.g. car, crane, copy machine), (4) semiautomatic tools (e.g. fax, milking installation, bottling machine) (5) and computer-based tools (e.g. computers, CNC machines). I adopt this categorization for all waves of the survey.

[46] BIBB: Federal Institute for Vocational Education and Training; IAB: Institute for Employment Research; BAuA: Federal Institute for Occupational Safety and Health.