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

Cäcilia Lipowski*

[Link to latest version]

February 29, 2024

Abstract

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

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

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

---

*ZEW Mannheim, caecilia.lipowski@zew.de

I am greatly indebted to Anna Salomons and Ulrich Zierahn-Weilage for their advice and support. For helpful comments and conversations, I thank Daron Acemoglu, Melanie Arntz, David Autor, Eduard Brüll, 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) 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

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

In this paper, I overcome this identification issue by exploiting an education reform and provide empirical evidence on the causal effect of supply-driven labor shortages on firm technology investments. I focus on shortages of young labor market entrants, i.e. trainees; a decisive, yet under-researched factor: First, young workers are currently becoming increasingly scarce due to population aging. Second, their availability may be the bottleneck to technology adoption if, as I will argue throughout the paper, new technologies require up-to-date skills and trainees excel in learning new skills compared to incumbent workers.

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

The German vocational training system provides an exceptional opportunity for studying implications of shortages of young labor market entrants for several reasons. First, vocational

[1] 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.

[2] 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).

training is omnipresent in the German labor market with two thirds of the workforce holding a vocational training degree.[3] Second, its institutionalized set-up allows for the precise identification of trainees, training firms, and training periods in administrative data. Third, the low geographic mobility of trainees greatly enhances the sharpness of the negative trainee supply shock with respect to state boundaries, aiding identification. Fourth, trainee wages are highly rigid because they are largely set by collective bargaining agreements. Consequently, labor supply shocks are unlikely to engender wage responses (as I also document in the case of the missing trainee cohort), leading to what is defines a labor shortage: below-equilibrium employment.

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

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 degree, henceforth "highly educated trainees". Highly educated trainees make up 16% of all trainees (Federal Statistical Office, Genesis-Online, 2022a), while the majority of trainees have 9 or 10 years of schooling, henceforth "low-educated trainees." Training wages do not increase. Firms also do not compensate missing highly educated trainees by hiring more low-educated trainees or workers with completed vocational training. Commuting of trainees across states does not intensify, and internal training of incumbent workers is not expanded.

A second key finding is that trainee shortages cause reductions in investments: investments decrease sharply in training firms in treated states compared to training firms in control states in face of the trainee shortage. This finding addresses the central question this paper raises: Trainees and investments are complements rather than substitutes, and their scarcity does not induce firms to invest more in order to compensate their absence but rather impedes investments. The effect is large: While highly educated trainees represent 3% of a training firm's workforce (11% of a training firm's hires), investments per worker temporarily drop by approximately 20% in affected years. In line with a literature emphasizing the lumpy nature of investments (e.g. Doms & Dunne, 1998; Cooper et al., 1999), the large average investment

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

decline consist of firms refraining from *large* investment projects, combined with firms not reducing investments because they did not plan to invest in the counterfactual scenario. I confirm the link between the investment decline and the absence of trainees in several ways. First, non-training firms in non-training industries do not reduce their investments. Further, firms compensating the lack of trainees by hiring non-trainees also reduce investments, indicating that the investment response is specific to the shortage of trainees rather than a general labor shortage. Next, in an auxiliary identification strategy, I show that firms more affected by the negative trainee supply shock decrease investments to a greater extent using a Bartik-type instrument exploiting pre-reform exposure to the shock. This finding also ensures that the investment decline in the event study design is not (exclusively) driven by firm selection into hiring trainees despite the shortage.

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

To rationalize the effects of the trainee shortage on technology investments, I present a stylized economic framework of endogenous technological change that incorporates technology vintages and capital adjustment costs of worker training in vintage-specific skills. Standard models of endogenous technological change are unable to explain both the relevance of young entrants compared to incumbent workers and the impact on technology adoption despite constant factor prices, i.e. wages. In my framework, young labor market entrants complement technology adoption due to their comparative advantage in the acquisition of vintage-specific skills: Opportunity costs of training in terms of foregone output are low for young, initially unproductive, labor market entrants, and concomitant productivity gains of training are large. When young labor market entrants are scarce, firms refrain from adopting new technologies because costs of retraining incumbent workers are prohibitively high.

I provide empirical evidence in support of the mechanism via adjustment costs of worker training. The hypothesized mechanism implies that firms facing more pronounced skill changes reduce technology adoption more when trainees are scarce. Based on Lipowski et al. (2024), who demonstrate that skill in vocational training curricula are updated due to technological change, I provide empirical evidence of this implication. The investments drop is also more pronounced in firms with higher trainee retention rates, i.e. firms employing trainees as an investment in skills for future production.

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

Second, I contribute to a literature on technology-induced new skills and new tasks, and their impact on workers of different ages. The literature provides many examples of how new technologies require new skills and create new tasks, without ruling out the replacement of existing tasks (e.g. Chari & Hopenhayn, 1991; Autor et al., 2003; Acemoglu & Restrepo, 2018; Autor et al., 2022). Such new skills have been linked to decreasing returns to experience, early retirement, and reduced hiring opportunities for older workers (e.g. Aubert et al., 2006; Ahituv & Zeira, 2011; Deming & Noray, 2020). In consequence, adaptation to technological change takes place through the entry of young workers, rather than by upskilling incumbent workers (e.g. MacDonald & Weisbach, 2004; Cavounidis & Lang, 2020). Adão et al. (2020) argue that this reasoning particularly applies to the ICT revolution studied in this paper because of its substantial demand for new, specific skills. My paper provides causal evidence of the reverse relationship, manifesting that technology-specific skills induce capital adjustment costs of worker training, and hinder the technology adoption when young labor market entrants are scarce.

Third, I contribute to a nascent literature on the consequences of labor shortages on firm outcomes. Existing studies establish a negative effect on firm capital, sales, and productivity (D'Acunto et al., 2020; Le Barbanchon et al., 2023; Sauvagnat & Schivardi, 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 for 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 and its consequences 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 for occupations that typically require bachelor's or associate degrees in other countries, such as the US.

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

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

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 working contract for the duration of their vocational training and a training wage, even though training wages are usually subject to collective bargaining agreements and are low.[5] 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 not firm-specific and current.

After completing the training usually lasting three years, a high share of trainees remain at their training company.

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 with approximately 5%.

Trainees play a key role in firms' acquisition of new tech skills, as a representative firm survey suggests: when asked about their vocational training, 44.5% of the firms agree that it ensures the constant supply of new skills, 46.5% agree that it improves the firms' innovative capacity, and 43% agree that it enhances the firms' adaptability to market and technological changes.[6] Similarly, Schultheiss & Backes-Gellner (2022) show that in Switzerland, a country with a vocational training system similar to Germany, changes in training curricula result in firms being closer to the technology frontier.

2.2 The reform

Prior to German reunification in 1990, upper track school graduates underwent 12 years of schooling in East Germany (Mecklenburg-Western Pomerania, Brandenburg, Saxony, Thuringia, Saxony-Anhalt, East Berlin) and 13 years in West Germany. After reunification, in an effort to align the the two education systems, Brandenburg switched to 13 years in 1994, while Saxony and Thuringia retained the 12-year system. 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. The education reform was decided in May 1996 in Mecklenburg-Western Pomerania and in January 1998 in Saxony-Anhalt.[7] By length-

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

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

[7] 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 2005.

ening 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? While two thirds of the missing upper track school graduates would usually opt for university studies, one third would subsequently start vocational training. The missing school graduates of spring 2001 are hence expected to result in a missing entrant 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 trainee supply by one year.

Based on the official training statistics, Figure 1, Panel B shows that the missing school graduation cohort indeed translates into a missing cohort of newly concluded training contracts in 2001 and 2002. The number of training contracts with highly educated trainees dropped by approximately 60% in 2001 and 2002 in treated states, and much less across control states.

The missing trainees should also be visible in the firm panel used for the subsequent analyses. This data captures trainee employment with a lag of one year, because it is based on records as of June 30 each year, and most trainees start on August 01 each year. Hence, I expect a missing trainee entrant cohort in the data in 2002 (2003 for those who would have done military service) and a substantially reduced stock of trainees in 2002, 2003 and 2004 (and a slightly lower stock in 2005 due to military service). Indeed, hires of highly educated trainees decline by approximately 50% in 2002 in treated firms, corresponding to 0.16 trainees per firm, see Panel C.[8] The stock of highly educated trainees is reduced by approximately one third in 2002–2004, corresponding to 0.34 trainees per firm, see panel D.[9]

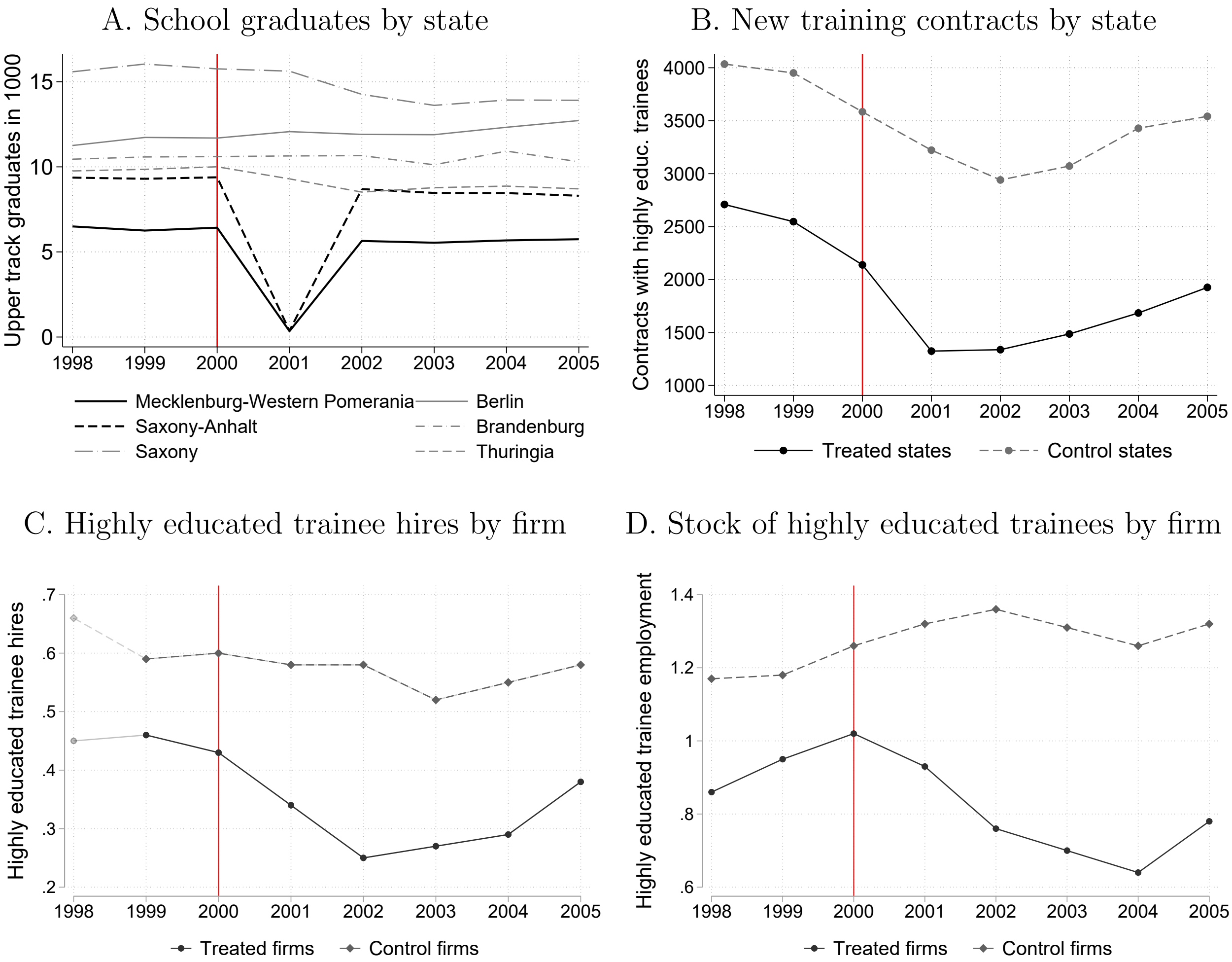
This is the labor supply shock in upper track vocational trainees I exploit for identification. I focus on upper track school graduates who subsequently start vocational training instead of university students/graduates because vocational trainees postpone their labor market entry less and move or commute less across federal states, thus endorsing the credibility of the identification strategy. Indeed, there is no visible decrease in the number of workers with tertiary education in affected states compared to control states, see Appendix Figure B1.1, Panel B,

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

[9]The lack of highly educated trainees might also translate into fewer workers with completed vocational training several years later. However, this shock is largely mitigated over time, see Appendix Figure B1.1, Panel A.

probably due to their high mobility across federal states and their tendency to postpone labor market entry.[10]

Figure 1: The missing school graduation cohort



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

In what follows, I assign Saxony-Anhalt and Mecklenburg-Western Pomerania as treated states and the other four East German states as control states. 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, in contrast, situated in central Germany with a population of around 2.2 million in 2020, features a comparatively more urban environment. Bordering Western Germany, Saxony-Anhalt is characterized by chem-

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

ical industry, mechanical engineering, and automotive supply. Both states are characterized by high unemployment rates during this time period, namely 17.8% in Mecklenburg-Western Pomerania and 20.2% in Saxony-Anhalt in 2000. Unemployment rate of the control East German states is not markedly lower (between 15.4 and 17.0%), see Appendix Figure B1.2, Panel B.

The education reform was a claim 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, lacking a majority in the government. In Saxony-Anhalt, the were the senior governing party but shared power with the Greens. Convincingly, 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 other policy or socio-economic changes confound the effect of the education reform, I compare several state metric including population size, education expenditure, unemployment rate, GDP, public debt and public investments between treated and controls 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, see Table 1. Controlling for state and year fixed effects, I find no difference in any of these outcomes between treated and controls states post-reform compared to pre-reform that is statistically significantly different from zero at the 10% level, see Panel A. Turning to factors correlated with the governance of the Social Democrats, see Panel B, there is significant positive association between government of the Social Democrats and three indicators: education expenditures, unemployment rate, and public investment. Higher education expenditure should, however, rather *in*crease instead of *de*crease firms technology investments. 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.

3 Firm panel data

Data sources. My analysis is based on the Linked-Employer-Employee-Data of the IAB[11] (LIAB), which combines the IAB Establishment Panel survey with administrative employment information of all employees at surveyed firms.[12] The IAB Establishment Panel is a large annual representative survey of establishments that includes information about investments, organizational change, sales, and internal training, among others. The Establishment Panel has existed in West Germany since 1993 and in East Germany since 1996. The number of surveyed establishments has risen from 4,000 in 1993 to 16,700 in 2020. Importantly, the survey

[11] IAB: Institute for Employment Research.

[12] 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.

Table 1: Correlation of state metrics with reform/government of the Social Democratic Party

	(1) Log(Population)	(2) Log(Education expenditure)	(3) % education expenditure	(4) Unemployment rate	(5) Log(GDP)	(6) Log(Public Debt)	(7) Log(Public Investments)
			A. Education reform in 2001				
Treated × 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	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 % relative to the dependent civilian labor force. Debt: Debt of the overall public budget. Mean: mean of dependent variable. 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) * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

is conducted at the workplace level, enabling the distinction between treated and untreated establishments based on their location.[13] I use the terms "firm" and "establishment" interchangeably for simplicity. Employment information is based on administrative records reported to the social security insurance. While employment information is reported as of June 30 each year, most vocational training programs start in fall, such that new trainees usually appear in the data with a lag of one year.

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

Data preparation. I restrict the dataset in four steps. First, I limit the data to firms in East Germany including Berlin, since the reform affects firms in East Germany which might not be perfectly comparable with firms in West Germany. Second, I exclude firms in the health/education/social services sectors because vocational training in many related occupations is purely school based. The reform therefore affect firms' trainee employment to a much lesser extent in these sectors. Third, I limit the sample to firms with at least ten employees each year, as larger firms usually possess more accurate data and more consistent behavior

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

over time. The results are robust to including smaller firms. Last, I constrain the sample to a balanced firm panel containing firms existing and with non-missing investments for the entire time period 1998–2005. A balanced panel has two main advantages over an unbalanced panel. First, it reduces compositional differences in the event study estimates that would likely violate the parallel trends assumption. Second, the firm-level matching procedure is only meaningful if treated and matched control firms are observed in the same years. As a disadvantage of a balanced panel, firms exiting the market or firms with missing values due to survey non-response are dropped. I find that firm exit is not affected by the reform, see Section 6. Conditioning on firm survival should hence not bias the estimates. With respect to survey non-response, I impute missing values by exploiting the panel dimension of the data. I proceed in two steps. First, I linearly interpolate missing values in up to two consecutive years if the firm has valid entries before and afterwards. This corresponds to imputing 2.1% of investment values but preserves an additional 13.0% of balanced firms. Second, I constantly extrapolate values at the start (1998, 1999) and at the end of the observation window (2004, 2005) for firms existing in these years as indicated in the social security records. At the cost of imputing 9.8% of investment values, this allows me to keep another 83.1% of balanced firms. Overall, by imputing 12% of investment values, the imputation procedure enables the inclusion of more than twice as many firms. For training firms, the share of imputed values is even lower. See Appendix A.1 for more details regarding the imputation procedure. The imputation procedure successfully recovers small firms with smaller investments which otherwise would have been lost due to the balancing requirement, enhancing the representativeness of the sample. I compute robustness checks which confirm the results in the non-imputed and/or unbalanced dataset.

Summary statistics. The final sample comprises 1,386 firms, of which 463 are treated (244 in Saxony-Anhalt and 219 in Mecklenburg-Western Pomerania) and 923 are untreated. Table 2 shows summary statistics of the final dataset. In sum, all firms cover approximately 3.5% of the East German workforce in a year.[14] I observe on average 11,396 trainees per year, of which 1,558 (13.7%) are highly educated, corresponding to 1.12 highly educated trainees per firm, or 0.61% of a firm's workforce. Based on the data, common occupations for highly educated trainees are media service occupations, retail occupations, insurance and financial service occupations, or technical drawer. Highly educated trainees are most common in the business service sector with an employment share of more than 2%, but can also be found in the manufacturing sector with an employment share of 0.3%, see Appendix Figure A2.1.

Training versus non-training firms. In 78% of the firm-by-year observations, no highly educated trainee is employed, and 59% of the firms never employ a highly educated trainee over the entire time window 1998–2005, see Table 2. Since the reform affects highly educated

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

Table 2: Summary statistics

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

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

Table 3: Pre-reform averages across training versus non-training firms

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

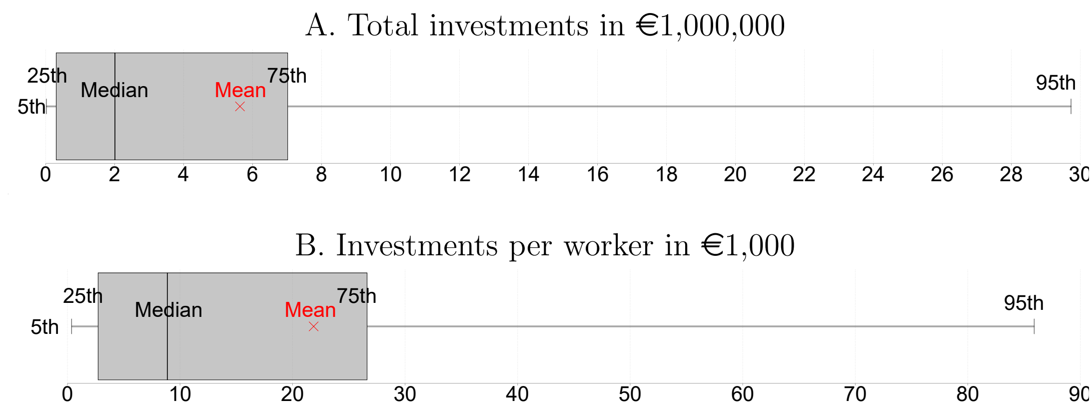
Notes: Average values across 1998, 1999 and 2000 of training and non-training firms. Δ: Average in non-training firm minus average in training firms. A training firm is defined as a firm with at least one highly educated trainee in 1998 and as non-training firm otherwise. Selected industries: industries with a significant difference between training and non-training firms. Hidden industries: Agriculture; energy, water, waste; retail/motor vehicles; transport; other services. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

trainee employment only, I focus on training firms, defined as firms with at least one highly educated trainee in 1998. This divides the sample into 293 training firms and 1,093 non-training firms. I base this classification on 1998 to minimize potential anticipation concerns. Less strict definitions of training firms, i.e. firms with at least one highly educated trainee in 1998 or 1999, or firms with at least on trainee in 1998 independent of the trainee's education show mitigated effects, as expected. Non-training firms in non-training industries are used in a falsification test. Table 3 shows summary statistics for training and non-training firms. In years prior to the reform, highly educated trainees made up 2.45% of a training firm's workforce and 10.76% of a training firm's hires. Training and non-training firms are fundamentally different. Compared to non-training firms, training firms are three times as large in employment, invest more, and operate more often in the business service and public administration sector, and less often in manufacturing and construction.

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, furnitures and fixtures, which I will term "production facilities", (2) communication technology, electronic data processing; "information and communication technologies (ICT)", (3) real estate and buildings; "real estate", and (4) means of transport, transportation systems; "transport". If a firm invested in at least one of these, the firm is surveyed on the total amount of annual capital investments. Accordingly, the investment volume is expected to contain investments in these four categories, while it is unlikely that non-tangible assets other than ICT are included. Although the data lacks details about the specific investment made or the specific technology adopted, it encompasses a broad spectrum of investments and technologies rather than concentrating solely on one. Appendix Table A2.1 provides a detailed description of the underlying survey questions and variable construction.

To curtail the influence of extremely large investments, I cap values in the top percentile of either total investments or investments per worker. The distribution of investments is highly right-skewed, see Figure 2, Panel A. While 19% of the observations show no investment, the median of all positive investments is approximately €2,000,000 per year, while the mean of all positive investments is close to €6,000,000 per year, with the 95th percentile at €30,000,000. The skewness might be attributed to the firm size distribution being right-skewed as well, or to the lumpy pattern of investments as emphasized in the literature (e.g. Bessen et al., 2023). To ensure that large firms do not drive the results, I focus on investments per worker, defined as total nominal investments divided by the initial number of workers in 1998, see Panel B.[15]. The median investment per worker and year is close to €10,000, the mean is €22,000, and the investment at the 95th percentile exceeds €80,000 per worker. Investments per worker are highest in the energy/water/waster sector, followed by public administration and business services; and lowest in construction and hospitality, see Appendix Figure A2.2, Panel A. On average, 35% of investments were attributed to firm expansion.

Figure 2: Distribution of firm investments



Notes: Observations at the firm-year level. Among training firms only.

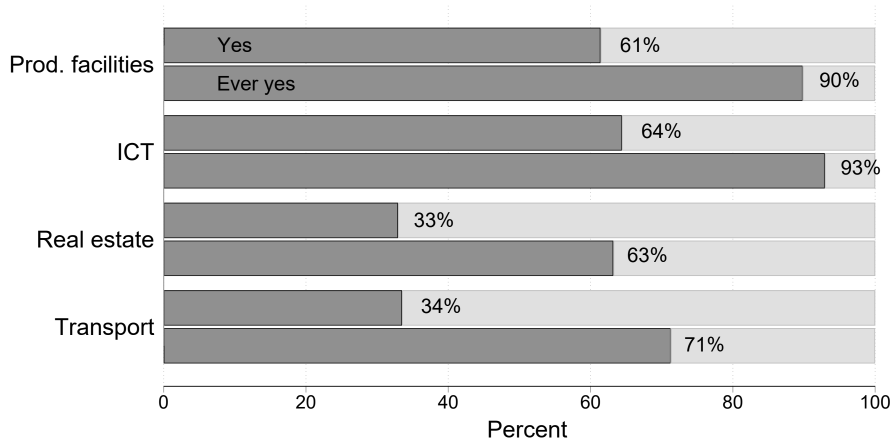
The establishment panel lacks a direct measure of the capital stock. To fill this gap, I exploit information on total investments, the proportion of net investments, dummy variables

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

representing four investment types, and the information on the industry. Equipped with these ingredients, I apply the modified perpetual inventory method developed by Müller (2008, 2017) explicitly for this dataset to estimate the capital stock. I establish a starting value for the capital stock using investments in the first three observed years (1996, 1997 and 1998 at the earliest). Then, leveraging this proxy, I project the capital stock for subsequent years using investment data and sector-specific depreciation rates. Please note that the capital stock therefore likely becomes more accurate over time. However, acknowledging the inherent inaccuracies in this method, I focus on investments while also reporting results for the log capital stock to assess the effect size.

Technological change. In order to determine if investments incorporate new technologies, I use three additional pieces of information from the data: investment types, the technical status of firms' machinery, and firm-level organizational changes, each of which I describe in the following. Figure 3 shows the share of firms investing in one of the four investment types per year and at least once over the time window 1998–2005. More than 90% of the firms invest in production facilities or ICT at least once between 1998 and 2005. In 61% of the firm-by-year observations, I observe an investment in production facilities, and in 64% an investment in ICT. These shares do not vary strongly by industry, see Appendix Figure A2.2, Panel B. Investments in real estate and transport occur less frequently with one third of firm-year observations making an investment in either category. 3.5% of all firms never invest between 1998 and 2005. This share is lower among training firms (2%). In summary, capital investments occur regularly in the data.

Figure 3: Investments by types



Notes: Observations at the firm-year level. Yes: Firm-year observations with investment in a certain investment type. Ever yes: Firms invested in a certain investment type in any year between 1998–2005.

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

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

Table 4: Technological change and investment types

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

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

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

4 Event study approach

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

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

where Y is one of several outcomes such as investments, j denotes the firm, b the federal state, and t the calendar year. Treated is a binary variable with Treated $= 1$ if the firm is located in a state undergoing the reform and zero otherwise. ψ_t capture calendar-year fixed effects. State fixed effects $\phi_{b(j)}$ capture time-constant level differences between federal states. The results are robust to including firm fixed effects. The vector β_t includes the coefficients of interest, namely the differential firm outcomes in treated states compared to firms in control states following the reform in 2001. The event study thus identifies the causal effect of a firm facing a state-wide negative trainee supply shock.[16] I stop in 2005 because of a different education reform affecting trainee supply from 2007/2008 onwards. Note that treatment is not staggered, precluding potential biases common to two-way fixed effects estimators in a staggered setting (e.g. Goodman-Bacon, 2021).

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

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

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

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

entire time window 1998 to 2005.[17]

Treated training firms may differ from control training firms in aspects which expose treated training firms to different potential confounders than control training firms. Indeed, treated training firms operate less often in manufacturing than control training firms. In terms of other conceivably relevant pre-reform firm characteristics, such as the share of highly educated trainees or investments, treated training firms and control training firms differ remarkably little, see Appendix Tables B1.1 and B1.2. To ensure that treated and control training firms are comparable, I match treated and control firms based on their pre-treatment characteristics. The matching procedure consists of two steps. In a first step, I match firms within training and non-training status and within nine industry groups. By matching within industries, the estimated reform effects are devoid of distorting industry-specific shocks or heterogeneity in treatment effects by industries. In a second step, I perform 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 and pre-treatment investments per worker. I directly match on investments in all pre-treatment periods since investments cannot be well approximated by other covariates due to their lumpiness (Bessen et al., 2020). Due to the potential issues with restricting analysis to firms with no pre-existing trends, as highlighted by Roth (2022), I present findings for both the entire sample of firms and a matched sample throughout the paper. In essence, Strategy 1, without matching, is preferable if one worries about anticipation, and Strategy 2, with matching, is more suitable if one worries about firm-specific shocks. Convincingly, the results are similar for both samples and robust to the matching specification.

The matching procedure does not provide remedy if external factors evolve differently in treated and control states. As shown in Section 2.2, major state metrics such as unemployment, population size, education expenditure and public debt and investments do not significantly change in treated compared to control states post 2001.[18] Moreover, one might be concerned that the introduction of the euro in 2002, the German Hartz reforms over 2003–2005, the bust of the dot-com bubble in 2000, or China's accession to the World Trade Organization in 2001/2002 might confound the reform effect. However, these shocks likely affected treated

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

[18] Zooming in on population growth and the unemployment rate, I observe very comparable patterns across states, see Appendix Figure B1.2. 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.

East German states and control East German states similarly, especially within industries. In addition, it is unclear why any other shock would affect firm outcomes differently based on the share of highly educated trainees at a firm.

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

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

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

There are two aspects that potentially changed due to the reform but the change was permanent, namely the skill level of highly educated trainees due to the increased years of schooling, and the share of upper track graduates starting vocational training. Effect dynamics will help distinguishing these permanent adjustments from the temporary trainee shortage.[19]

One might be concerned that the missing trainee entry cohort entails concomitant consumer demand changes, as labor supply changes due to migration usually entail. In the present case, however, consumer demand is unlikely to adjust since the overall population size remains

[19] Also, higher levels of education would, if any, likely induce more investments, and therefore provide a lower bound of the effect.

constant, and per capita spending likely adjusts only marginally given trainees' low wages. Turning to firm demand, low trainee wages also prevent a meaningful decrease in the firm wage bill, 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.[20] 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 absence of the shock,[21] and $\hat{\beta}_t$ are large if trainees are primarily missing in firms that would have invested in 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: 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 to identify a different causal parameter, namely the effect of employing one trainee less.

Inference. The standard advice is to cluster standard errors at the level of treatment assignment to account for cluster-level shocks (e.g. Abadie et al., 2023). In this setting, the number of clusters, i.e. federal states, is small. For valid inference with a small number of clusters, I follow three approaches. First, as suggested by Roth et al. (2023), I assume that the cluster-specific shocks are small compared to the idiosyncratic error terms at the firm level, potentially resulting in a small violation of parallel trends. The remaining uncertainty comes from the sampling of firms within clusters only. I hence cluster standard errors at the firm level. This assumption is well justified in the data: The variance of the error term across all observations is approximately equal to the average variance within states but much smaller within firms, suggesting little to no within-state correlations but large within-firm correlations. Second, I report confidence intervals based wild t-bootstraps clustered at the state level as suggested by Cameron et al. (2008). Third, I perform permutation (Fisher randomization) tests, comparing the t-statistic of the treatment effect for the actual treatment assignment and for all permuted treatment assignments across federal states.

[20] 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.

[21] For example, the investment decline would be zero, if only never-investors forego trainee employment. This is an unlikely scenario, since only 2% of the training firms never invest over the observed time period.

5 Bite of the reform

Figure 4 displays the results of the difference-in-differences event study model outlined in equation (1) regarding the effect of the reform on trainee hires (Panel A) and trainee employment (Panel B). The left panel shows the result based on the sample of all firms, the right panel based on the sample of matched firms. Panel A shows a clear drop in hires of highly educated trainee in treated training firms compared to control training firms in 2002. This is precisely the year the majority of the upper track school graduates of 2001 would have appeared in the data as new hires.[22] With 0.64 fewer hires in 2002 in the sample of all firms (-1.11 in the sample of matched firms), this corresponds to a pronounced drop of approximately 30% (50%). Hires in treated training firms remain slightly below hires in control training firms in 2003 and 2004, likely due to postponed entry related to military service. The gap has closed by 2005. Pre-trends in hires should be interpreted with caution due to the challenging identification of hires in the dataset in 1998.[23]

Panel B focuses on the stock of highly educated trainees, which experiences a longer-term decline since vocational training usually takes three years. In 2002, 2003 and 2004, approximately 1.5 fewer highly educated trainees work in treated training firms compared to control training firms. With an average of 4.9 highly educated trainees per training firm in 1998, this corresponds to a drop by one third. Considering the typical training duration of three years, this aligns with the absence of one year's worth of upper track school graduates. Consistent with the timeline of the shock, the employment gap starts to shrink in 2005.[24] Firms' highly educated trainee employment evolves in parallel in control and treated states in the years 1998 to 2000, likely because there are no suitably qualified trainees available to employ in anticipation of the reform. Trainee employment already starts to drop in 2001, potentially due to some trainees already being employed at their training firms on June 30 before the official training start on August 1st. Convincingly, the estimates are comparable across the sample of all training firms and the sample of matched training firms.

Wage effects. The data allows to study firms' adaptation strategies, such as wage changes or the substitution of highly educated trainees with other workers. To investigate such accompanying effects, I employ the corresponding difference-in-differences specification, comparing the pre-treatment period 1998–2000 to the post-treatment period 2002–2004. Results are given in Table 5.

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

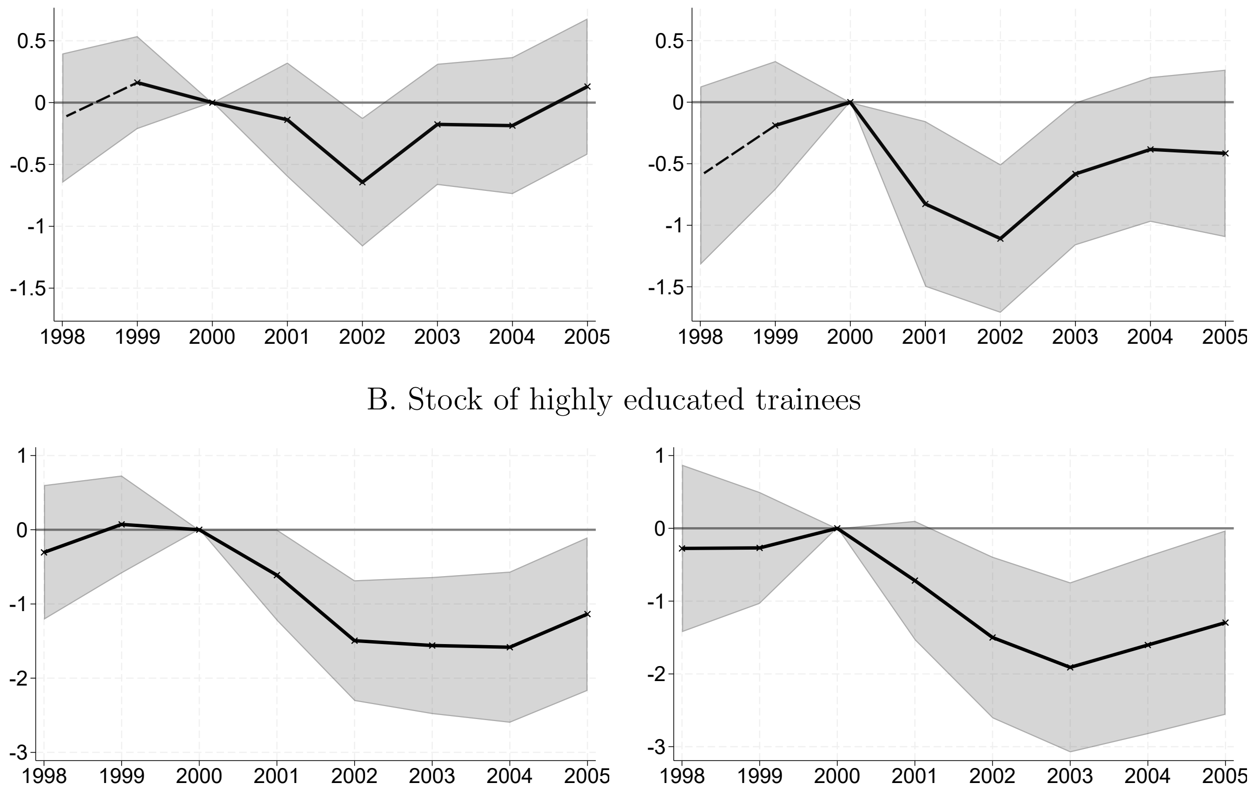
[23] Since firms do not report new hires themselves, I impute hires based on observed individual employment, exploiting the panel dimension of the dataset. For firms entering the panel in 1998 it is hence impossible to determine whether an employee is a new hire or an incumbent.

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

Figure 4: Effect on trainee employment

A. Highly educated trainee hires

All training firms | Matched training firms



Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Standard errors clustered at the firm level. Hirings in 1998 should be taken with caution. Training firms only. Number of observations: Panel A – left: 2,295; Panel A – right: 1,233; Panel B – left: 2,344; Panel B – right: 1,248. For the corresponding graph with confidence intervals based on cluster wild t-bootstraps, see Figure B2.1.

In contrast to what standard economic theory predicts, there is no evidence of a wage increase among highly educated trainees in response to the negative supply shock (column 1). This result is not caused by firm selection into training highly educated trainees in the years of reduced trainee supply, as the specification with firm fixed effects reveals (column 2). This finding is in line with the results by Muehlemann et al. (2022) in the case of the opposing, positive supply shock of trainees. To understand the absence of the wage effects, it is important to keep in mind that supply of school graduates is fixed by the cohort size. Therefore, the purpose of raising trainee wages would be confined to either poaching trainees from other firms, attracting non-school graduates, e.g. the unemployed or already employed workers, or attracting school graduates from other states. The lack of such attempts can be attributed to factors linked to both the temporary nature of the shock and features of the German system: First, firms likely shy away from raising 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 hypothetical doubling of training wages would result in negligible changes in absolute income. Instead, trainee supply responds to anticipated post-training wages (Neuber-Pohl et al., 2023) that remain unchanged. Third, the vast majority of training wages are set by collective bargaining agreements,[25] with a single standard wage set for each region and industry, irrespective of occupation or school education, making wage adjustments unlikely. Of course, firms could deviate upwards. In that case, works councils, which would have to approve training wages in large firms, would likely oppose unequal treatment of trainees.

Table 5: DiD Results – Wage and worker substitution effects

	Wage effects		Substitution effects			
	Log wages highly educ. (1)	trainees (2)	# low-educ. trainee hires (3)	# highly educ. commuting trainee hires (4)	Internal retraining (5)	Log VT employment (6)
	Empirical strategy 1 – All training firms					
Treated × Post	−0.03	−0.04	−0.62	0.03	−0.09*	−0.13**
	(0.04)	(0.03)	(0.85)	(0.05)	(0.05)	(0.05)
	[-0.11;0.08]	[.]	[-3.13;1.44]	[-0.04;0.09]	[-0.25;0.08]	[-0.24;-0.07]
N	1758	1758	2295	2018	2227	2344
Firm FE		X				
Init. outcome	3.00	3.00	6.81	0.04	0.42	4.87
	Empirical strategy 2 – Matched training firms					
Treated × Post	0.03	0.01	−0.54	0.09	−0.12	−0.12*
	(0.05)	(0.04)	(1.07)	(0.07)	(0.08)	(0.07)
	[-0.07;0.14]	[.]	[-2.09;1.62]	[-0.08;0.35]	[-0.27;-0.02]	[-0.25;-0.03]
N	908	908	1233	1082	1190	1248
Firm FE		X				
Init. outcome	3.00	3.00	6.51	0.03	0.43	4.82

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Post: 2002–2004. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 90% confidence bands based on cluster wild t-bootstraps in square brackets. Init. outcome: Average outcome of treated firms in 1998. Column 4: a commuter is defined as a person living and working in two distinct federal states. This variable is available from 1999 onward. Column 5: Internal retraining is the sum of retraining incidences at the firm-year level. VT: completed vocational training. For the full set of results, see Appendix Table B2.1. For further outcomes see Appendix Table B2.2.

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

[25] In 2022, at least 82% of trainees were covered by collective bargaining agreements (Schönfeld & Wenzelmann, 2023) and even firms that are not part of those agreements tend to base their wages on such agreements.

related to distinct skill sets, the specialization in different occupations, and the stable demand for low-educated trainees.

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

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

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

6 Effects on firm technology investments

6.1 Overall effect on investments

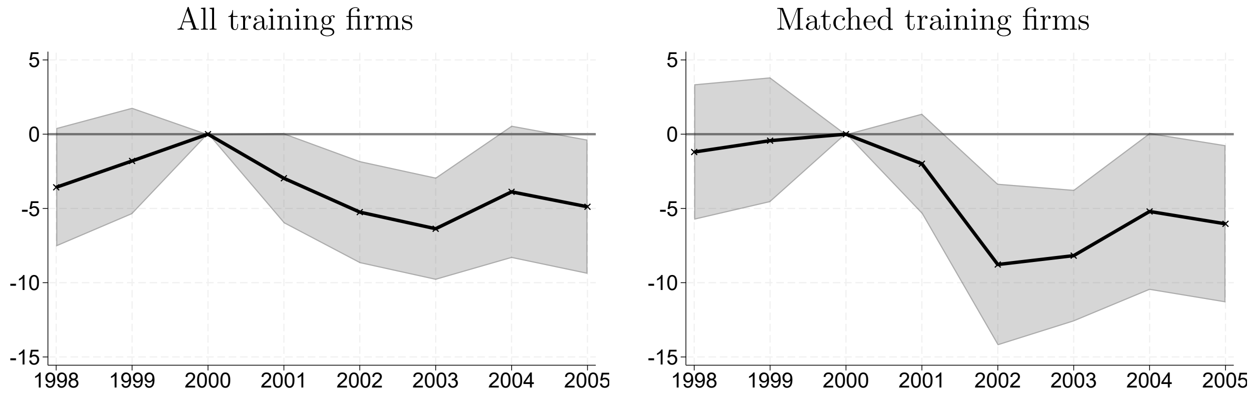
I now turn to the impact of the negative trainee supply shock on firm investments. Figure 5 shows a large and statistically significant decline in investments per worker following the reform in treated training firms compared to control training firms; the key finding of this paper. The strongest investment decline is observed in the years 2002 and 2003. Across the three years of missing trainees, 2002, 2003 and 2004, this corresponds to a drop of 19% of the average value in 1998 (sample of all training firms; 37% in the sample of matched training firms).

There are no statistically significant pre-trends.[26] No anticipation in investments is consistent with the idea sketched out in Section 7 that new technologies arrive constantly and firms are unable to adopt them once trainees are missing. The negative effect diminishes after 2003, corroborating its relation to the temporary drop in trainee supply. While it seems that investments do not fully return to their initial level by 2005, this finding is not statis-

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

tically significantly different from zero and not robust across specifications. Importantly, the reform-induced foregone investments are not recouped at least until 2005. This implies that even a temporary trainee supply shortage leads to a permanent reduction in the capital *stock* (assuming no immediate capital depreciation).

Figure 5: Effect on investments per worker in €1,000



Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Standard errors clustered at the firm level. Outcome: investments in €1,000 divided by total employment in 1998. Training firms only. Number of observations: Left: 2,344; Right: 1,248. For the corresponding difference-in-differences estimates, see Table 6. For the corresponding graph with cluster wild t-bootstrap confidence intervals, see Figure B2.2.

I employ alternative specifications of the outcome variable in corresponding difference-in-differences regressions. The results are given in Table 6. First, I ensure that the effect is not driven by changes in the denominator, i.e changes in employment. When dividing investments by the current number of workers instead of the initial number of workers (column 2), I find a negative but smaller effect on investments per worker, indicating that the investment decline is partly but not exclusively driven by a concomitant employment decrease.

Next, I account for the highly right-skewed distribution of investments in combination with frequently observed zeros. Since a simple log-transformation has been acknowledged to be problematic (e.g. Chen & Roth, 2023), I instead employ several alternative transformations suggested by Chen & Roth (2023). In particular, I separate the effect on the extensive margin from the effect on the intensive margin (columns 3 and 4). This allows to log-transform investments for strictly positive values. Investments decline only insignificantly and little at the extensive margin (3–5 percentage points, corresponding to 5–6%), but substantially at the log-transformed intensive margin (33 – 38%). In order to combine both margins, I manually assign an importance to the extensive margin. In particular, I define a change from zero to any strictly positive investment to be as important as an investment increase by 1% (column 5). This combined measure indicates an investment drop of 7–10% of its initial value. Compared with other estimates in the literature, the estimated decrease corresponds approximately to the decline the literature would predict if capital costs permanently increased by 9-15% (Zwick

& Mahon, 2017; Lerche, 2022; Liu & Mao, 2019).[27] To classify the scale of the effect, it is useful to consider the effect on the log capital stock. While the value is approximated only, it suggests that the capital stock decreased by 7–8 log points (column 7).

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

Table 6: DiD Results – Investment effects

	Investments per worker		Intensive vs. extensive margin				
	per init. # of workers	per current # of workers	Any inv. (0/1)	Log(Inv.)	Combined	Large inv. (1/0)	Log(K)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			Empirical strategy 1 – All training firms				
Treated × Post	-3.36^{*}	-1.00	-0.05	-0.33^{**}	-0.64^{***}	-0.07	-0.08
	(1.85)	(2.49)	(0.03)	(0.16)	(0.24)	(0.04)	(0.05)
	[-6.91;-0.06]	[-2.94;0.98]	[-0.14;-0.00]	[-0.53;-0.18]	[-1.17;-0.37]	[-0.09;-0.03]	[-0.34;0.13]
% of init. outcome	-19%	-6%	-6%		-10%	-12%	
N	2344	2344	2344	2069	2344	2069	2271
Init. outcome	17.68	17.67	0.89	7.46	6.63	0.56	10.31
			Empirical strategy 2 – Matched training firms				
Treated × Post	-6.55^{***}	-4.20	-0.03	-0.38^{*}	-0.48	-0.06	-0.07
	(2.36)	(3.48)	(0.05)	(0.22)	(0.37)	(0.07)	(0.15)
	[-10.23;-3.21]	[-8.24;0.83]	[-0.16;0.11]	[-0.68;-0.04]	[-1.36;0.48]	[-0.16;0.07]	[-0.34;0.13]
% of init. outcome	-37%	-23%	-3%		-5%	-11%	
N	1248	1248	1248	1102	1248	1102	1191
Init. outcome	17.93	17.93	0.88	7.41	6.55	0.57	10.26

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Post: 2002–2004. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Combined: Explicitly combining the extensive and intensive investment margin by assigning a change at the intensive margin to be as important as an investment increase by 1%. Large inv.: Investments in the upper tercile of the investment per worker distribution assigned as one, and zero otherwise. Log(K): Log of the imputed capital stock. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 90% confidence bands based on cluster wild t-bootstraps in square brackets. Init. outcome: Average outcome of treated firms in 1998. For the full set of results, see Appendix Table B2.3. For further outcomes see Appendix Table B2.4.

Effect size. The average investment decline is large despite the fact that highly educated trainees make up only 2.5% of a training firm's workforce. In particular, the investment decline goes beyond a potential "mechanical" effect of reducing capital in proportion to trainee employment.[28] Why do trainees so heavily impact firm investments? The answer is manifold. First, as the stylized economic framework in Section 7 will detail, trainees play a crucial role

[27] Zwick & Mahon (2017) study bonus depreciation in the US between 2001 and 2010, finding that a 1% reduction in investment costs increases investments by 3.69 log points. Lerche (2022) estimates an increase in investments by 2.43 log points in response to a 1% reduction in investment costs in the setting of investment tax credits in East Germany in 1999 among manufacturing firms. Liu & Mao (2019) find a value of 2.26 in China.

[28] In addition, 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 €487 on average per year

in the adoption of technologies requiring new skills. These technologies are likely 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, if exposed. 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. To assess this hypothesis empirically, I plot the distribution of the underlying firm-level matched difference-in-differences estimates in Figure 6. Indeed, in line with this reasoning, the average treatment effect is the combination of firms not reducing their investments and firms foregoing large investments. To explicitly analyze the effect on large investments, I run a difference-in-differences regression among observations with strictly positive investments using a binary outcome taking the value one for investments in the upper tercile of the investment per worker distribution (>€10,000), and zero otherwise, see Table 6, column 6. Treated training firms are 6–7 percentage points (11–12%) less likely to make large investments than control training firms when trainees are scarce. The effect is comparable when focusing on investments per worker in the upper decile (>€51,200), see Appendix Table B2.4, column 5, and is even more pronounced when defining large investments within industries, see Appendix Table B2.4, column 6.

Second, Figure 6 also reveals that the distribution of the investment drop is highly right-skewed, leading to excessively large average treatment effects compared to, for example, median treatment effects: The average investment drop in 2002 – the statistic reported throughout the paper – is four times larger than the median investment drop.[29]

Third, the trainee shortage is temporary and known to be so. There is no reason for firms to invest in exactly the period when trainees are short in supply, leading them to postpone investments. the absence of a rebound in investment once the trainee shortage diminishes is perplexing and could be linked to firms unexpectedly put on different trends, or firms skipping over certain technology vintages.

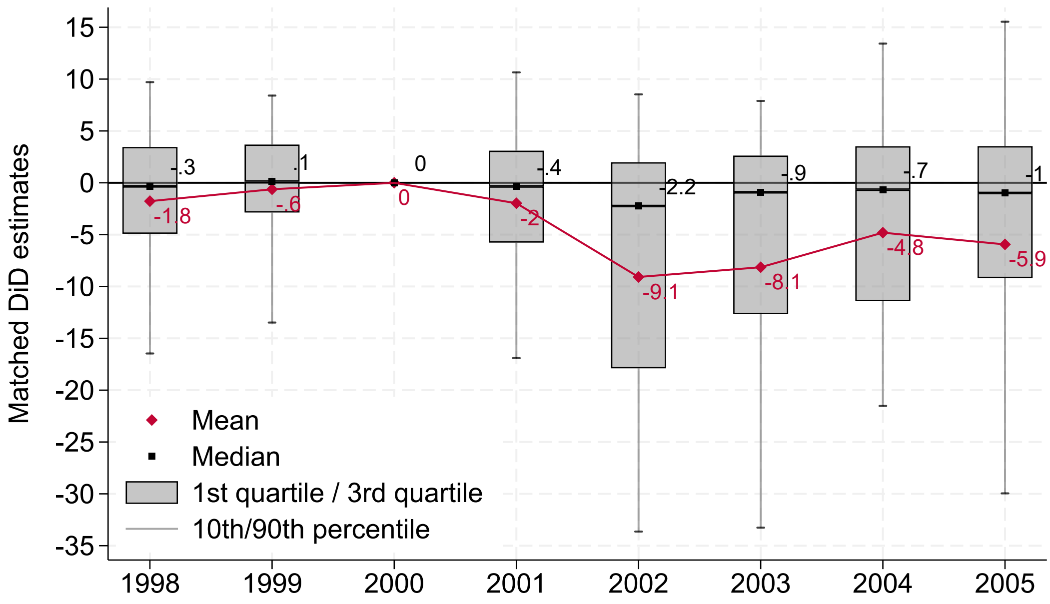
Also, trainees might not only directly affect firm investments, but also via additional indirect channels. For example, trainees might impede firm employment growth, reducing firm investments further. While this is likely to be a relevant margin inflating the treatment effect, this is not the only reason behind reduced firm investments, as I show below.

---

and trainee on equipment and material (Beicht et al., 2004). With a reform-induced reduction in the number of trainees by 1.50 in 2002 and an average size of training firms of 354 workers this would imply a "mechanical" reduction of €2.06 per worker. In addition to these €487 on equipment and material costs, East German firms in 2000 reported €1530 of "other costs" per trainee per year, including costs for teaching material, fees, and training administration. If a firm interpreted all these costs as capital investments, the total "mechanical" reduction in investments would still be as small as €8.55 per worker.

[29] Outliers in the uppest percentile of the total investment distribution or in the uppest percentile of the investment per worker distribution are trimmed. Conducting quantile regressions is not feasible due a lack of sufficient observations for this more demanding model.

Figure 6: Distribution of matched DiD estimates – Investments per worker in €1,000



Notes: Distribution of the matched firm-level difference-in-differences estimates. Outcome: Investments per worker in €1,000. *Red:* Average. *Black:* Median. *Box:* 25th and 75th percentile. *Adjacent values:* 10th and 90th percentile.

Falsification test among non-training firms. Next, I turn to the sample of non-training firms. Non-training firms should not be directly affected by the reform. Confirming this hypothesis, the average investment drop among non-training firms is less than half of the average drop among training firms, see Table 7, columns 1 and 2. See Appendix Figure B2.3 for the event study results. The investment drop among training firms is, however, not exactly zero. This could be due to either a poor approximation of non-training firms, indirect effects such as industry spillovers, or, in the worst case, confounding factors. To address the first point, I expand the set of non-training firms from those without a highly educated trainee in 1998 to those without a highly educated trainee in the entire pre-period 1998 until 2000, see column (4). This does not remove the negative, partly insignificant effect among non-training firms and might still include firms that would have liked to hire trainees in the years of the shock. The negative effect among non-training firms might also stem from indirect effects such as industry spill-overs related to product market competition, knowledge spill-overs, or poaching of workers with completed vocational training. To control for such indirect effects within industries, I perform a difference-in-differences regression including the triple interaction term between Treated, Post, and the share of highly educated trainees in an industry in 1998, while controlling for all corresponding two- and one-way interaction terms (Table 7, columns 3, and 5). This analysis indeed reveals that there is no negative treatment effect among non-training firms in non-training industries.

Table 7: Falsification test and industry spillover

	Training firms	Non-training firms			
		Strict definition		Broad definition	
	(1)	(2)	(3)	(4)	(5)
		Empirical strategy 1 – All firms			
Treated × Post	−3.37*	−1.53	0.87	−1.53	1.53
	(1.79)	(1.21)	(3.58)	(1.26)	(3.73)
Treated × Post × Industry exposure			−2.40		−3.15
			(3.95)		(4.18)
N	2344	8744	8744	8024	8024
		Empirical strategy 2 – Matched firms			
Treated × Post	−6.86***	−2.73*	0.04	−2.92**	1.18
	(2.29)	(1.47)	(5.00)	(1.35)	(4.11)
Treated × Post × Industry exposure			−2.76		−4.14
			(5.57)		(4.55)
N	1248	6112	6112	5616	5616

Notes: Outcome: investments per worker in €1,000. Industry exposure: Industry share of highly educated trainees in 1998 in %. *Strict definition:* Firms without any highly educated trainee in 1998. *Broad definition:* Firms without any highly educated trainee in 1998-2000. Reference group: Treated × Pre. Pre: 1998–2000. Post: 2002–2004. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Standard errors clustered at the firm level. See Appendix Figure B2.3 for the event study results.

Robustness. The negative effect of the reform on investments per worker is robust to a large range of data samples and specifications. To see this, I present difference-in-differences estimates in Figure 7, comparing the post-reform years 2002–2004 with the pre-reform years 1998–2000. I show the estimates for both the set of all training firms and the set of matched training firms. The coefficients are consistently larger in magnitude and statistically more significantly different from zero in the matched sample.

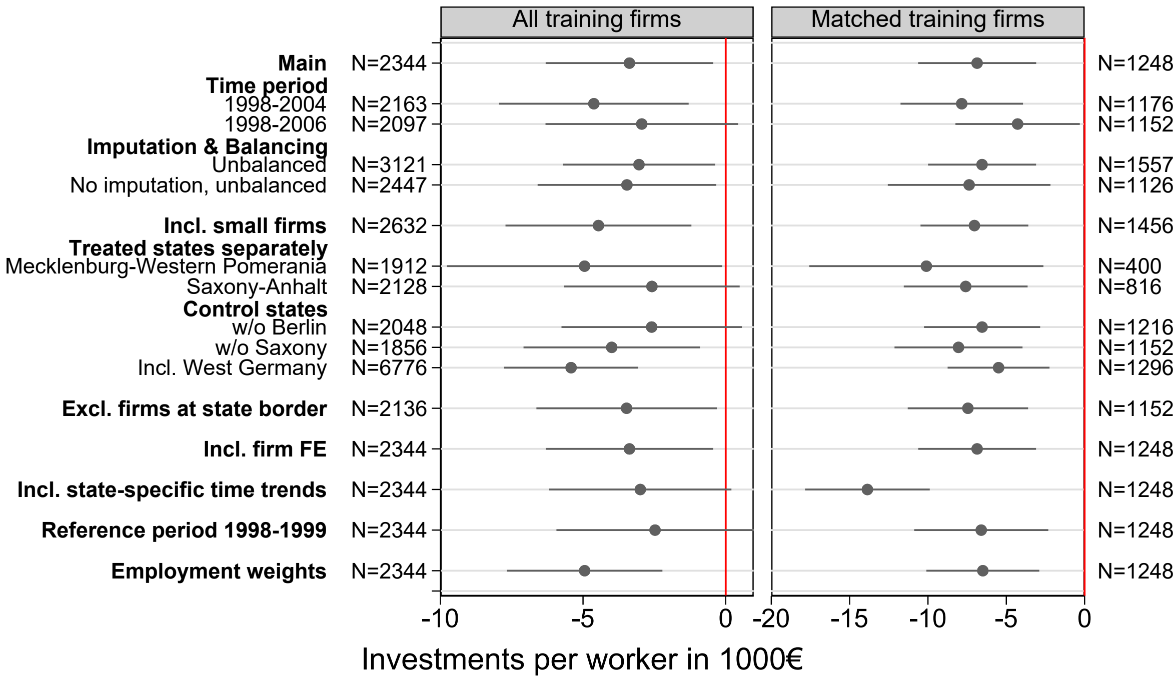
The negative estimate is robust to the specification of the balancing requirement, i.e. when restricting to firms observed for the entire time period 1998 to 2004, or 1998 to 2006 instead of 1998 to 2005, or when fully abolishing the balancing requirement. The result is also virtually unchanged when not imputing missing values.

Next, I confirm that including firms with less than 10 employees does not alter the results significantly either. Convincingly, the effect is found within both treated states separately, despite their differences in industry structure and geography. When excluding Berlin or Saxony-Anhalt from the set of control states due to their slightly different demographic and economic trends, the result remains robust. Expanding the set of control firms to include West German firms substantially increases sample size but produces similarly sized negative estimates as well.

Firms at federal state borders might be less affected by the reforms because they may attract trainees from control states. Since firm address is not disclosed in the data and counties are too large too 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 higher commuter share does not affect the results meaningfully.

The coefficient is not affected by the inclusion of firm fixed effects instead of state fixed effects.[30] Allowing states to be on different (linear) time trends does not change the size of the estimate in the set of all training firms and only further increases the estimate in the set of matched training firms. I next exclude the year 2000 from the pre-period because it might be distorted by anticipation effects. Again, the size of the estimate is virtually unchanged, but deceases in precision. Last, I weighting the observations by the firms' initial employment size in 1998. This increases the negative coefficient, indicating that the impact per individual is more pronounced than the impact per firm.

Figure 7: Robustness



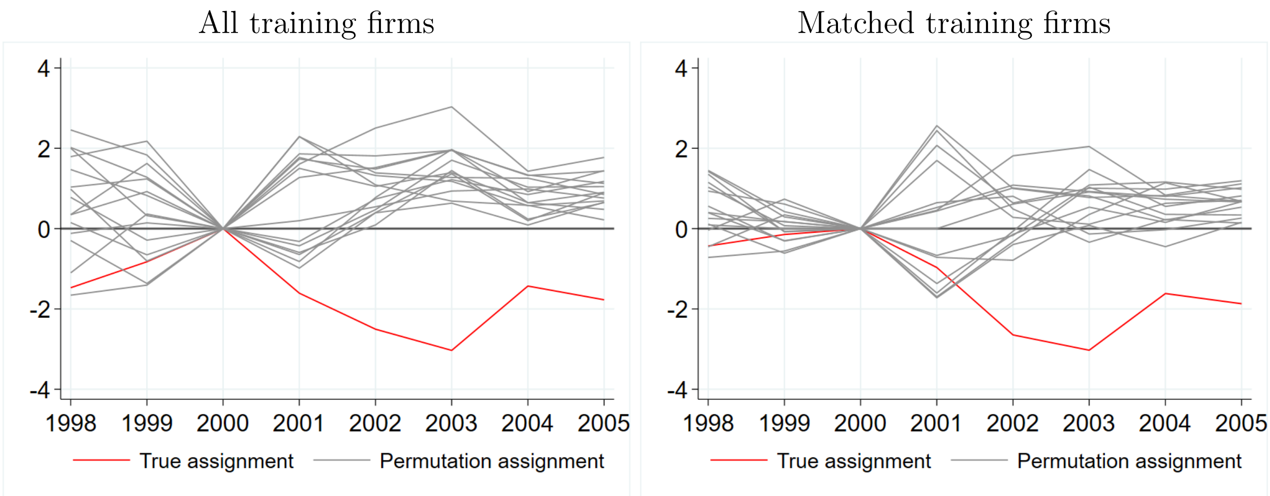
Notes: DiD estimates of the term Treated × Post (2002–2004) compared to Treated × Pre (1998–2000) plus 90% confidence bands. Standard errors clustered at the firm level. N = number of observations. *Main:* Main specification. *Time period:* Balanced firm panel for 1998–2004 and 1998–2006. *Unbalanced:* Including firms with missing investment values. *No imputation, unbalanced:* No imputation + including firms with missing investment values. *Incl. small firms:* Including firms with less than 10 employees. *Treated states separately:* Only using firms from one treated state and dropping firms from the other treated state. *Control states:* Dropping Berlin or Saxony from the set of control firms. Additionally including all West German training firms as control firms. *Excl. firms at state border:* Dropping those 10% firms with the highest cross-state commuter share of workers with vocational training in 1999. *Incl. firm FE:* Including firm fixed effects instead of federal state fixed effects. *Incl. state-specific time trends:* Additionally controlling for linear state-specific time trends. *Reference period 1998–1999:* Excluding the year 2000 from the pre-period. *Employment weights:* Observations weighted by firms' initial employment size in 1998.

[30] Since treatment status does not vary across states and all firms are observed in all calendar years, the two estimates are equivalent.

Permutation tests for inference with few clusters. Until now, I have assumed that the cluster-specific shocks are small compared to the idiosyncratic error terms at the firm level, justifying the use of standard errors clustered at the firm level. I next perform permutation tests which have been suggested a valid method for inference when the number of clusters is small (e.g. Roth et al., 2023). Figure 8 shows the t-statistics for the event study estimates based on the actual treatment assignment in red, and for all permuted treatment assignments across East German federal states in gray. The t-statistics are based on standard errors clustered at the firm level and account for sampling error of firms within states. Following the reform in 2001, the t-statistics based on the actual treatment assignment are by far more negative than any t-statistic based on a permuted treatment assignment. For periods prior to the reform, this is not the case, suggesting no differential pre-trends. Hence, the permutation test shows that it is very unlikely that cluster-level shocks only would have caused the observe investment decline. Likewise, no comparable decrease in employment of highly educated trainees was observed under any permutation assignment, see Appendix Figure B2.4, Panel A. This result hold for both the sample of all firms and the sample of matched firms.

Since the number of possible permutations within East Germany is limited to 15, I repeat the permutation test across the 10 West German federal states. There was no comparable education reform in West Germany around that time. The t-statistics of the uppest and lowest 5% of the draws under permuted treatment assignment are shown in Appendix Figure B2.4, Panel B. Again, the t-statistic of actual treatment assignment stands out as an outlier much smaller than the 5% and 2.5 most negative t-statistics under permuted treatment assignment.

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



Notes: T-statistics of the event study coefficients of the actual treatment assignment (red line) and permutation assignments within East Germany (gray lines). Outcome: investments in €1,000 divided by total employment in 1998. For additional results, see Appendix Figure B2.4.

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

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

6.2 Effect on firm technology adoption

Having established that the reform-induced trainee shortage decreases overall capital investments, the following section investigates whether this decrease is linked to foregone technology adoption by studying the effect on direct indicators of firm-level technological change. Results based on the equivalent difference-in-differences estimation as above are given in Table 8. As a first measure of firm-level technological change, I look at the self-assess technical status of a firms' machinery on a scale from 1 ('completely out-of-date.') to 5 ('state-of-the-art') (column 1). Unlike investments, technical status is a *stock* variable, expected to deteriorate as foregone investments accumulate. I therefore focus on the year 2005, when missing investments of the years 2002–2004 have accumulated. Treated training firms report an outdated technical status of their machinery compared to control training firms in 2005. The depreciation is meaningful in magnitude and statistically significant, at least for the set of all training firms (Panel A): the coefficient of -0.18 is equivalent to 18% of the firms reporting a reduction by one category. Since the reported technical status only changes from one year to the next only in 30% of the observation, this corresponds to half of all firm-level technological changes. Confirming the link to the trainee shortage, the falsification test among non-training firms confirms that there is no depreciation of the technical status of machinery in non-training treated firms compared to non-training control firms, see Appendix Table B2.6. The result is comparable in magnitude but statistically not significantly different from zero for the sample of matched firms.

Table 8: DiD Results – Effects on firm-level technological change

	Technical status	Organizational change	Investment type (0/1) Production facilities	ICT	Real estate	Transport *(Placebo)*
	(1)	(2)	(3)	(4)	(5)	(6)
			Empirical strategy 1 – All training firms			
Treated × Post	-0.18^{**}	-0.37^{**}	-0.07	-0.08^{*}	-0.07	-0.01
	(0.09)	(0.16)	(0.05)	(0.04)	(0.04)	(0.05)
	[-0.31;0.04]	[-0.60;-0.22]	[-0.12;0.01]	[-0.11;-0.05]	[-0.15;0.01]	[-0.09;0.09]
N	2341	1311	2344	2344	2344	2344
Init. outcome	3.97	1.35	0.70	0.79	0.58	0.33
			Empirical strategy 2 – Matched training firms			
Treated × Post	-0.14	-0.66^{***}	-0.09	0.00	0.00	-0.02
	(0.13)	(0.22)	(0.07)	(0.06)	(0.07)	(0.07)
	[-0.31;0.12]	[-1.03;-0.28]	[-0.19;0.06]	[-0.05;0.08]	[-0.09;0.09]	[-0.14;0.14]
N	1245	702	1248	1248	1248	1248
Init. outcome	3.98	1.41	0.69	0.78	0.56	0.32

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Post: 2002–2004. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Technical status: Technical status of a firm's machinery on a scale from 1 ('completely out-of-date.') to 5 ('state-of-the-art'). Organizational change: On a scale from 0 to 4 as the sum of up to four organizational measures. ICT: Information & Communication Technologies. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 90% confidence bands based on cluster wild t-bootstraps in square brackets. Init. outcome: Average outcome of treated firms in 1998. For the full set of results, see Appendix Table B2.5. For the corresponding table based on non-training firms, see Appendix Table B2.6.

As a second direct indicator of firm-level technological change, I study firm-level organizational change (column 2). 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 following the reform. This decrease amounts to 0.37 (0.66 for the matched sample, respectively) reorganization measures less per firm, a drop by approximately one third (one half). Again, there is no comparable decline for treated non-training firms.

I next turn to investment dummies by investment types, see columns 3–6. Unfortunately, these measures only capture the extensive margin, while we know that the intensive margin is more heavily affected by the negative trainee supply shock. Treated training firms are 7–9 percentage points less likely to invest in production facilities. Although this effect is not statistically significantly different from zero, it is convincing that there is no effect among non-training firms. The probability to invest in in ICT is 8 percentage points lower in treated training firms compared to untreated training firms in the unmatched firm sample. For the matched sample, the effect on investments in ICT fades to zero. Investments in real estate, which are positively correlated to firm-level technological change, also are less likely when trainees are scarce, with this effect being not statistically different from zero. The probability to invest in transport remains unaffected. This is reassuring since we know from Section 3 that investments in transport are unrelated to technological change. I conclude that at least part of the investment decline is the result of reduced technology adoption. Foregone technological and organizational change, in turn, may explain the reduction in internal retraining of incumbent workers established above.

Foregone investments and a slow-down in technology adoption should affect firm performance in the longer-run. However, panel attrition and confounding shocks in later years pose problems when studying longer-term outcomes. I therefore look at firm performance indicators until 2005 only, see Appendix Table B2.7. Regarding sales per worker – a variable particularly poorly filled in the data – I find suggestive evidence for a decrease. Firm shrink in employment. It is unclear whether trainee shortages depress firm employment directly or whether foregone investments lead to employment declines. I find no effect on average log wages, or the probability of firm exits.

7 Stylized economic framework and supporting evidence

7.1 Stylized economic framework

I next present a stylized economic framework to rationalize the complementarity between young labor market entrants and technology adoption. A more detailed formalization is available in Appendix D. I build up on the endogenous technological change model in Acemoglu (1998) that highlights that technology adoption is endogenous to factor prices: If labor is scarce, increased wages incentivise the introduction of (skilled) labor-saving technologies and the reduction of (skilled) labor-complementing technologies. In the context of this paper, this standard model needs refinement for two reasons. First, labor is not generally scarce; the scarcity is limited to young labor market entrants. Second, factor prices, i.e. wages, do not adjust. To account for these two aspects, I introduce technology vintages and capital adjustment costs of worker training: Young labor market entrants have a comparative advantage in learning new skills required to handle new technology vintages due to lower opportunity costs of training and higher productivity increases. In consequence, their absence increases capital adjustment costs and hinders the adoption of technologies even when incumbent workers are abundant and wages do not adjust.

Consider the following set-up: Firms maximize profits by deciding whether to adopt a new, exogenously arriving and productivity-enhancing technology. Note that this decision is discrete. A new technology might substitute or complement labor in existent tasks. Crucially, the new technology always introduces at least one new task that requires skills specific to its vintage. In consequence, firms incur capital adjustment costs in terms of worker training if they want to adopt the new technology vintage.[31]

Firms can acquire skills either by retraining incumbent workers or by training young, initially unskilled, labor market entrants within a vocational training program.[32] Training costs consist of foregone production output during training and are incurred by the firms. Without training, production output of young labor market entrants is low, while incumbent workers are productive even without retraining. In consequence, firms prefer to train young labor market entrants over retraining incumbent workers because their opportunity costs of training are lower and their productivity increase due to training higher.[33] Note that the prediction is independent on whether only the opportunity costs channel, the productivity increase channel, or both apply.

When young labor market entrants become temporarily unavailable, firms adopt a new technology only if the productivity gain associated with the new technology is large enough to offset the increase in capital adjustment costs due to 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.

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

Alternative channels. There are two alternative explanations for the complementarity between young labor market entrants and technology adoption other than their low opportunity costs and great productivity gains of learning new skills. First, according to standard human capital theory, human capital investments in young workers yield longer-term benefits in expectation (the "horizon" channel in Cavounidis & Lang, 2020). Second, young workers might generally posses more up-to-date tech skills. While both channels may play a role, they cannot fully cause the observed investment decline because they 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 aspect new labor market entrants are considerably different to second-year trainees is in their opportunity costs and expected payoff of acquiring new skills, as noted in Cavounidis & Lang (2020). Indeed, the Cost-Benefit Surveys of Vo-

[31] Note that this process of skill acquirement might take place within occupations but might also require changes in the firms occupational composition.

[32] In principle, firms could also acquire these skills by poaching workers that have already acquired the new skills from other firm. 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.

[33] 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.

cational Training show that firm revenues from skilled labor activities of second-year trainees (third-year trainees) are 134% (254%) higher than of first-year trainees (Schönfeld et al., 2016, Table 18).

The reasoning above describes a potential mechanism that rests on two assumptions: New technologies require new skills, and trainees stay at their training firm. In the next section, I provide empirical evidence in support of both assumptions.

7.2 Empirical evidence on mechanism

New skills. The literature provides many examples of how new technologies require new skills, without ruling out the replacement of labor in existent tasks (e.g. Autor et al., 2003; Acemoglu & Restrepo, 2018; Deming & Noray, 2020; Autor et al., 2022). If the necessity of vintage-specific technology skills is the reason underlying firms' investment reductions as outlined above, firms that are more exposed to skill changes 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 changes depend on young labor market entrants to invest in new technologies because their incumbent workers do not possess the adequate skills. I measure occupational skill changes using changes in vocational training curricula from Lipowski et al. (2024). Training curricula offer an ideal approximation of skill changes for three reasons. First, they directly apply to the studied worker group, i.e. trainees. Second, their changes are caused by 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 1998 share of workers in occupations whose training curricula are updated around the time of the reform.[34] There is substantial variation in firm exposure to new skills with the median firm having 38% of their workers in changing occupations, the firm at the 10th percentile 6% and the firm at the 90th percentile 96%.

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

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

where j denotes a treated firm and j' its matched control firm. I regress this firm-level

[34] In the main specification I use curriculum changes between 1996 and 1999. Results are independent of the exact time period of curriculum changes considered.

difference-in-differences on firm exposure to new skills.

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

This approach is equivalent to a triple difference-in-differences specification with the triple interaction term Treated × Post × NewSkills plus all corresponding two-way and one-way interaction effects. In contrast, the specification in equation (3) is easier to interpret and allows to flexibly control for potential confounders X. In particular, I control for industry and firm exposure to the shock, i.e. number of highly educated trainees in 1998. I hence 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.[35]

The predicted investment change for firms with strong skill change (at the 90th percentile of the distribution) versus firms with little skill change (at the 10th percentile of the distribution) is shown in Figure 9, Panel A. In line with the hypothesized mechanism of capital adjustment costs of worker training in new skills, there is a negative association between firm exposure to new skills and firm investment drop, i.e. the predicted investment drop is larger among firms with stronger skill changes. This finding, besides imprecisely estimated, is similar when looking at curricula that are changed directly before the missing trainee cohort and when looking at curricula that are changed directly following the missing trainee cohort, see Appendix Table B3.1. This shows that the investment drop is not merely a direct effect of a new curriculum, but rather the result of a general skill change in an occupation which is, among others, expressed in a curriculum change.

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

I use the same strategy as in equation (3), but with the trainee retention rate as the independent variable of interest. The retention rate is defined as the proportion of trainees staying

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

Figure 9: Treatment effect heterogeneity

Notes: Predicted change in investments per worker in €1,000 in treated training firms compared to their matched control training firms between 2002 and 2000 as defined in equation (2). Prediction and 90% confidence bands based on equation (3) using heteroscedasticity-robust standard errors. Skill change: 1998 share of workers in occupations with an updated curriculum between 2000 and 2001. Trainee retention: Pre-reform share of trainees retained by the firm upon completion of the training. Little change/low retention: 10th percentile of the corresponding distribution. Strong change/high retention: 90th percentile. Controlling for industry and firm exposure to the reform (number of highly educated trainees in 1998). For the corresponding regression coefficients, see Appendix Tables B3.1, B3.2 and B3.3.

at the firm upon training completion.[36] Figure 9, Panel B, shows the predicted investment changes for firms with high and low trainee retention rates. Consistent with the hypothesized mechanism, firms with high retention rates reduce investments heavily in response to the reform, while treated firms with low retention rates reduce their investments much less. For the regression table, see Appendix Table B3.2.[37]

Ruling out the alternative mechanism via firm growth. These two heterogeneity analyses empirically support the hypothesized mechanism via the need for trainees to accompany the technology adoption due to their comparative advantage in skill acquisition. A shortage of young labor market entrants may also decrease firm investments because it creates an impediment to firm employment growth. If the impediment to firm employment growth is the only reason for the investment cut, only treated firms that indeed experience a net reduction in their workforce ("shrinking"), i.e. firms that do not substitute the missing trainees with other

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

[37] Since the economic framework predicts a U-shaped pattern of investment decline by the retention rate, see Appendix D, I also include the quadratic term in a further check. While the coefficients are not statistically significantly different from zero, they have the expected sign, see Appendix Table B3.2, columns 5 and 7.

workers, should reduce investments. In contrast, treated firms that replace the missing trainees ("not shrinking") should not reduce investments. I define shrinking firms as firms with a zero or negative absolute employment growth between 2000 and 2002, and as not shrinking otherwise. Figure 9, Panel C, shows that investments decline similarly strongly for shrinking and not shrinking firms. See Appendix Table B3.3 for the corresponding regression table. This finding is incompatible with the alternative mechanism via barriers to firm employment growth. It also demonstrates the unique role of young labor market entrants for firm investments: Firms may replace young labor market entrants with other workers, but these other workers are not able to fill the gap when it comes to firm technology adoption.

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.

This finding is informative from two perspectives. First, it highlights that the availability of young workers is a key factor determining firm technology adoption. Second, it informs on the detrimental effects that shortages of young workers have on firm investments. Assuming that even those technologies that generally replace labor require some new skills, the finding challenges hopes of addressing labor shortages by substituting labor with capital (e.g. Acemoglu & Restrepo, 2018). It hereby contributes an additional dimension to macro studies predicting economic downturn in times of population aging (e.g. Jones, 2022; Kotschy & Bloom, 2023; Maestas et al., 2023).

The proposed mechanism behind the complementarity between young workers and technology adoption, namely the comparative advantage of young workers in learning new skills due to lower opportunity costs and/or higher expected payoffs, is likely to hold in a broad range of settings. However, the relevance of this mechanism compared to other potential adjustment mechanisms depends on the context, in particular on the type of technology to be adopted, the type of scarce labor, the functioning of the labor market, and the time horizon of the labor shortage.

First, regarding the technology type, the more productive a technology, and the fewer new skills required, the more likely it will be implemented despite a labor shortage. Second, the type of scarce labor and its comparative advantage in skill acquisition will impact the relevance of the mechanism. While young workers have lower opportunity costs of training than incumbents in most conceivable setting, 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 labor supply shock can be absorbed by the labor market in different ways. If, for example, wages adjust such that employment of young workers does not decrease, the effect on technology adoption will be much different. Last, the effect of a temporary supply shock likely differs from the effect of a long-term reduction: Incentives to adopt labor-saving technologies, the channel highlighted in standard endogenous technological change models, are higher in the case of a long-term supply reduction.

In conclusion, while a reduction in the supply of young workers may not always causes a decrease in technology investments due to several other potential adjustment channels, it will always entail an increase in the costs of technology adoption.

From a policy perspective, the findings not only stress the importance of expanding measures to attract young labor market entrants, they also call for subsidies for retraining experienced workers. The results also have implications for the German vocational training system: While it seems to effectively foster the adoption of new technologies, as suggested 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., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. American Economic Review, 108(6), 1488–1542.

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

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

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

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

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

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

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

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

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

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

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

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

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

BIBB-Cost-Benefit-Survey 2000. (2009). doi: 10.7803/370.00.1.2.10

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Dorner, M., & Görlitz, K. (2020). Training, wages and a missing school graduation cohort. IAB-Discussion Paper, 28/2020.

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, jahre, 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

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

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

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

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

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

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.

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

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

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

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

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

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

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

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

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

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

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

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.

OECD. (2023). Retaining talent at all ages. doi: 10.1787/00dbdd06-en

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

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

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

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

Schönfeld, G., & Wenzelmann, F. (2023). Tarifliche ausbildungsvergütungen 2022: Erstmals im durchschnitt über 1.000 euro–anstieg aber unterhalb der inflationsrate. Fachbeiträge im Internet, BIBB.

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

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

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

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

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

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

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

A Data

A.1 Data imputation

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

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

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

Table A1.1: Data imputation steps

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

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

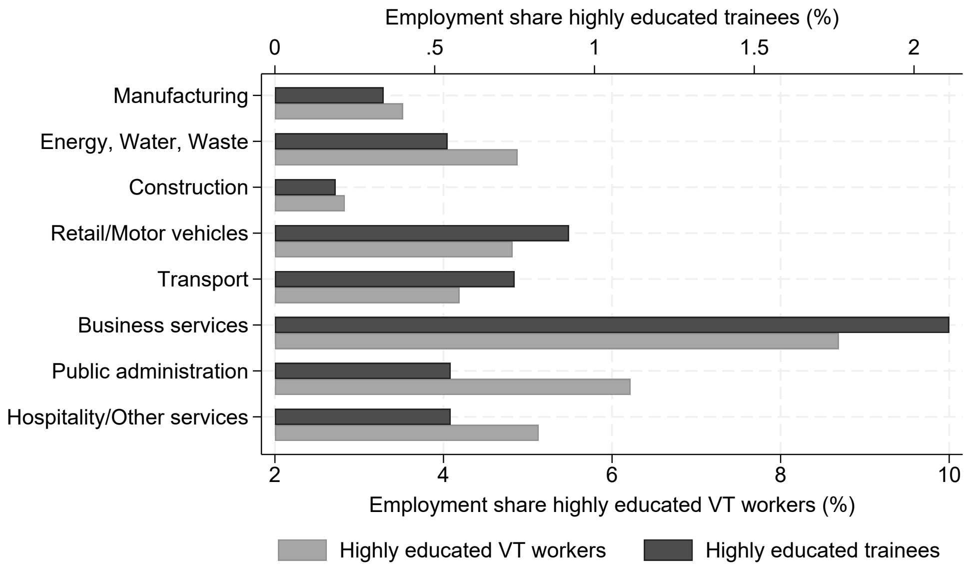
Table A1.2: Descriptives – Imputation and balancing

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

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

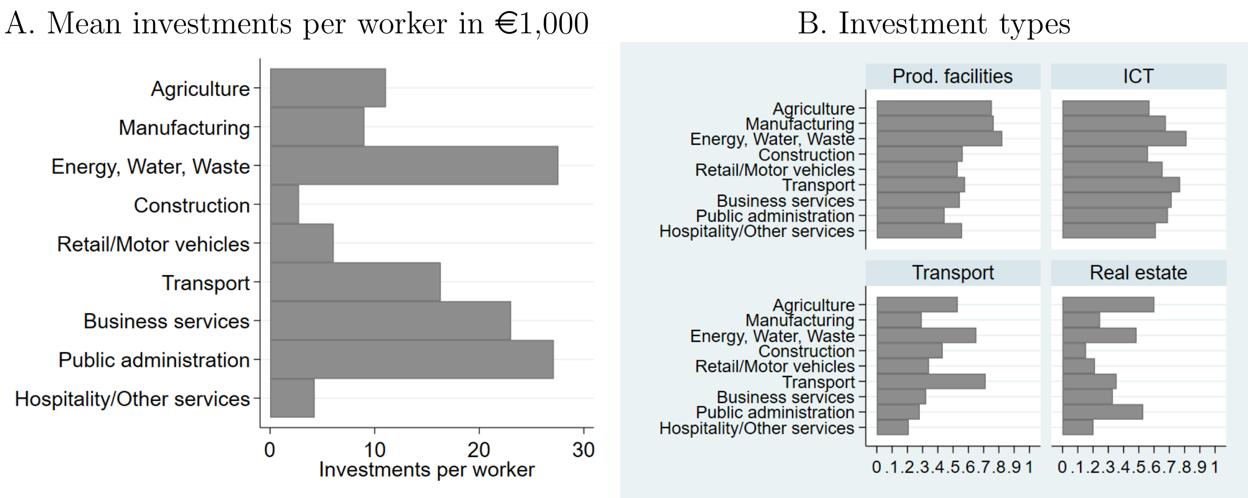
A.2 Descriptives and summary statistics

Figure A2.1: Highly educated trainee employment by industry



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

Figure A2.2: Investments by industry



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

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

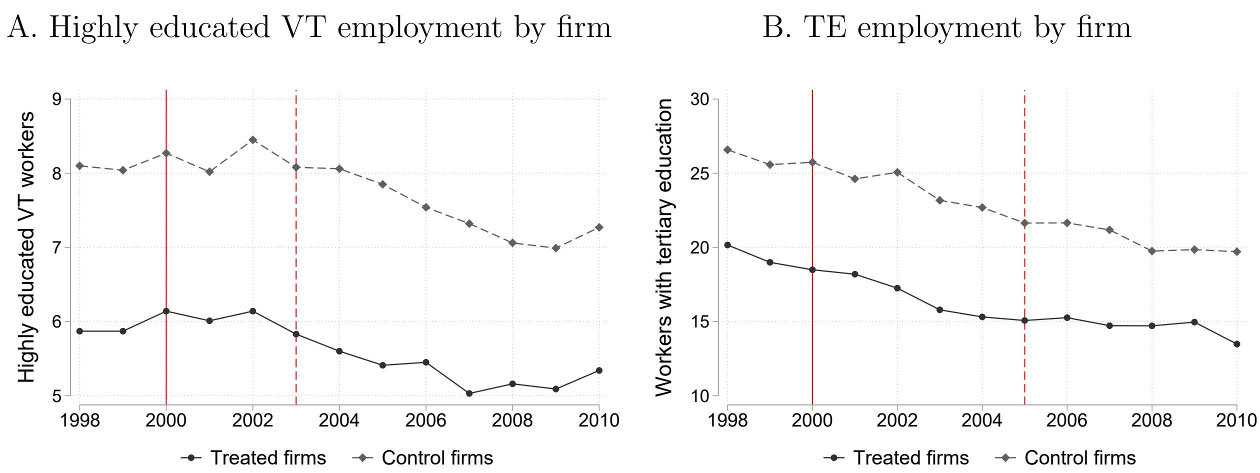
Variable	Survey Question	Manipulation	Frequency
Inv. per worker	What was the approximate sum of all investments in t?	Divided by number of workers in 1998 from the administrative records. Trimming uppest percentile of the investment distribution and the investment per worker distribution	Yearly
Expansion inv. per worker	What share of total investments made was attributed to the expansion of your establishment?		Yearly
Inv. in ICT	Did your establishment invest in one or more of the following areas in the last business year of t? EDP, information and communication technology?		Yearly
Inv. in production facilities	Did your establishment invest in one or more of the following areas in the last business year of t? Production facilities, plant and equipment, furniture and fixture?		Yearly
Inv. in transport	Did your establishment invest in one or more of the following areas in the last business year of t? Means of transport, transportation systems?		Yearly
Inv. in real estate	Did your establishment invest in one or more of the following areas in the last business year of t? Real estate and buildings?		Yearly
Technical status of machinery	How do you assess the overall technical status of the plant and machinery, furnitures and fixtures of this establishment compared to other establishments in the same industry? "1" - state-of-the-art equipment. "5" - completely out-of-date.	Inverted order	Yearly except for 2004
Organizational change	Has one or more of the following 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, 2010, 2012, 2014, 2015, 2017

Notes: t: Year of the survey. ICT: Information and communication technologies. EDP: Electronic data processing.

B Additional results

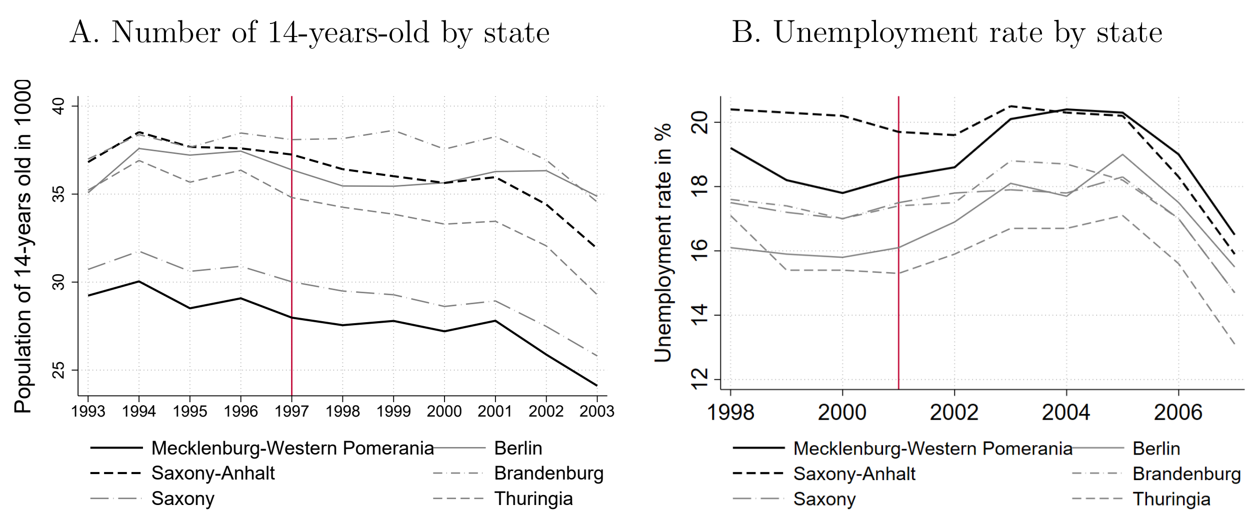
B.1 Additional results – Identification strategy

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



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

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



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

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

	Training firms			Non-training firms		
	Mean Treated	Δ Mean Unmatched	Δ Mean Matched	Mean Treated	Δ Mean Unmatched	Δ Mean Matched
	Matching within industries					
Agriculture	0.04	0.03	0	0.06	0.00	0
		(1.49)	0		(0.15)	0
Manufacturing	0.21	-0.11*	0	0.26	-0.10***	0
		(-1.91)	0		(-3.36)	0
Energy, Water, Waste	0.05	0.01	0	0.03	0.01	0
		(0.25)	0		(0.71)	0
Construction	0.05	-0.01	0	0.12	0.01	0
		(-0.41)	0		(0.70)	0
Retail/Motor vehicles	0.06	-0.03	0	0.11	0.03*	0
		(-0.84)	0		(1.69)	0
Transport	0.02	-0.01	0	0.04	0.01	0
		(-0.39)	0		(0.58)	0
Business services	0.21	0.03	0	0.12	0.01	0
		(0.49)	0		(0.47)	0
Public administration	0.31	0.08	0	0.18	0.03	0
		(1.31)	0		(1.16)	0
Hospitality/Other services	0.05	0.01	0	0.08	-0.00	0
		(0.42)	0		(-0.18)	0
	Mahalanobis distance matching					
% highly educated trainees 2000	2.32	0.31	0.43	0.12	-0.02	-0.01
		(0.73)	(0.83)		(-0.40)	(-0.29)
% highly educated trainees 1999	2.51	0.21	0.38	0.04	-0.01	-0.00
		(0.47)	(0.69)		(-0.38)	(-0.11)
% highly educated trainees 1998	2.68	0.05	0.28	0.00	0.00	0.00
		(0.11)	(0.49)			
Investment per worker 2000	18.59	0.58	3.77	13.96	-0.19	2.39
		(0.15)	(0.92)		(-0.09)	(1.03)
Investment per worker 1999	17.99	-1.33	3.14	14.66	-0.71	2.18
		(-0.37)	(0.85)		(-0.35)	(0.97)
Investment per worker 1998	17.68	-3.47	2.00	15.21	-0.17	2.03
		(-0.88)	(0.48)		(-0.08)	(0.85)
Pre avg. log(employment)	5.16	0.00	-0.19*	4.12	0.07	0.02
		(0.02)	(-1.13)		(1.13)	(0.26)
N		293	156		1093	764

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

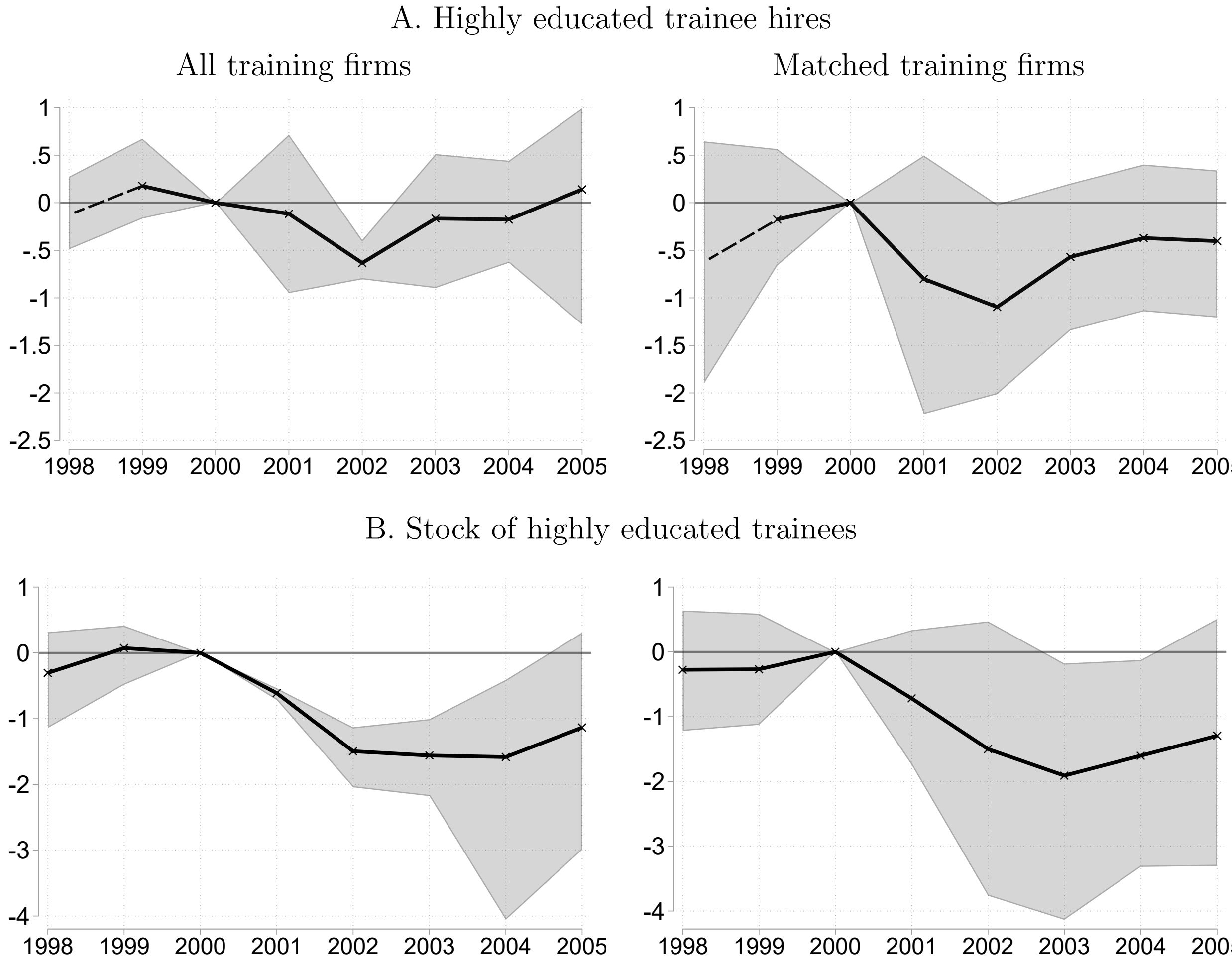
Table B1.2: Imbalance before and after matching – Non-targeted variables

	Training firms			Non-training firms		
	Mean Treated	Δ Mean Unmatched	Δ Mean Matched	Mean Treated	Δ Mean Unmatched	Δ Mean Matched
Any investment	0.92	-0.01	0.00	0.83	-0.04**	-0.02
		(-0.19)	(0.11)		(-2.21)	(-0.92)
Large investment	0.41	0.07	0.08	0.26	0.00	0.05
		(1.23)	(1.15)		(0.11)	(1.40)
Inv. in prod facilities	0.73	-0.06	-0.03	0.60	-0.06**	-0.04
		(-1.30)	(-0.55)		(-2.23)	(-1.44)
Inv. in ICT	0.85	0.02	0.00	0.64	-0.06	-0.04
		(0.52)	(0.00)		(-1.58)	(-1.16)
Inv. in real estate	0.56	0.05	-0.05	0.33	-0.01	0.03
		(0.92)	(-0.75)		(-0.45)	(1.04)
Inv. in transport	0.35	-0.07	-0.14**	0.37	0.01	0.05*
		(-1.42)	(-2.10)		(0.45)	(1.79)
Organizational change	1.20	0.13	0.20	0.68	-0.05	-0.03
		(0.92)	(1.23)		(-1.00)	(-0.48)
Technical status	3.98	0.05	-0.06	3.78	-0.09**	-0.04
		(0.59)	(-0.52)		(-2.10)	(-0.78)
Trainee retention rate	0.61	0.04	0.02	0.51	-0.05**	-0.05*
		(1.24)	(0.58)		(-2.30)	(-1.83)
Rate of skill change	11.36	-2.63	-0.81	16.03	0.23	1.19
		(-1.18)	(-0.31)		(-0.15)	(0.70)
N		293	156		1093	764

Notes: Numbers refer to the average in the pre-treatment period 1998–2000. Δ Mean: Mean Treated - Mean Control; N: Number of firms. T-statistic of the two-sided t-test of the difference in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

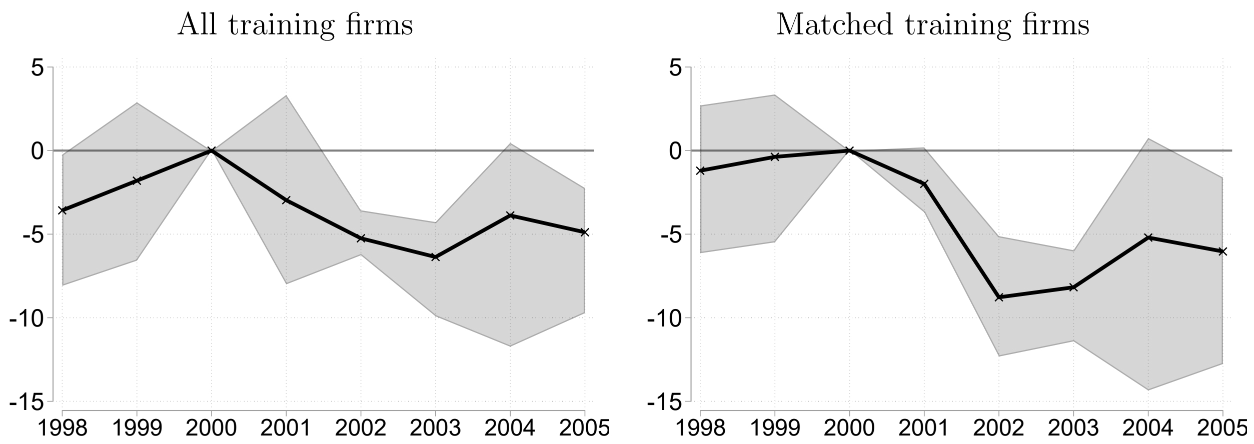
B.2 Additional results – Estimation results

Figure B2.1: Effect on trainee employment – Cluster wild t-bootstrap confidence intervals



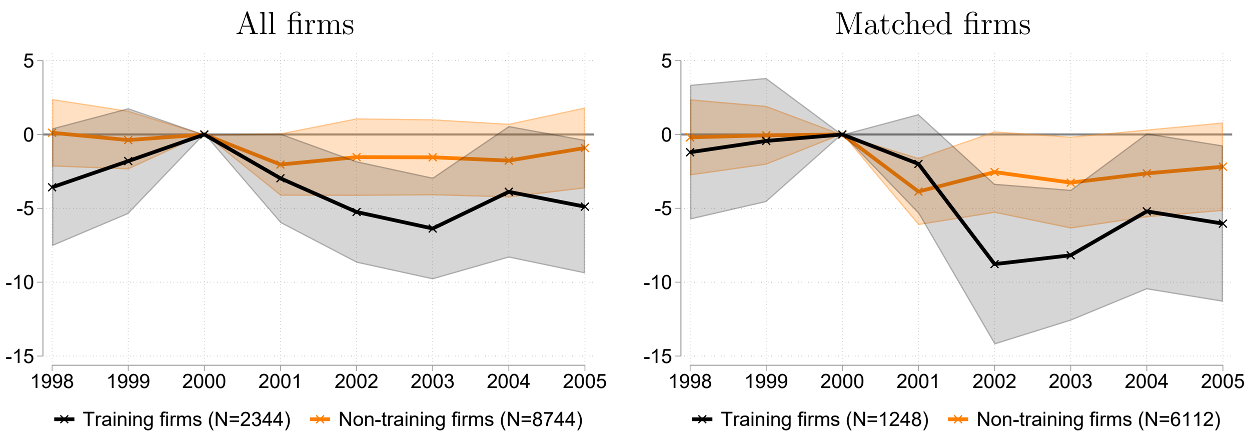
Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. 90% confidence intervals based on cluster wild t-bootstraps following Cameron et al. (2008). Hirings in 1998 should be taken with caution. For the main figure, see Figure 4.

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



Notes: Event study coefficients of the interaction terms Treated × Year plus 90% confidence bands. Outcome: investments in €1,000 divided by total employment in 1998. 90% Confidence intervals based on cluster wild t-bootstraps following Cameron et al. (2008). For the main figure, see Figure 5.

Figure B2.3: Effect on investments per worker in €1,000 in non-training firms



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

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

	Wage effects		Substitution effects			
	Log wages highly educ. trainees		# low-educ. trainee hires	# highly educ. commuting trainee hires	Internal retraining	Log VT employment
	(1)	(2)	(3)	(4)	(5)	(6)
			Empirical strategy 1 – All training firms			
Treated × Roll-out	−0.02	−0.05*	−0.04	0.05	−0.01	−0.05*
	(0.04)	(0.03)	(0.71)	(0.06)	(0.05)	(0.03)
Treated × Post	−0.03	−0.04	−0.62	0.03	−0.09*	−0.13**
	(0.04)	(0.03)	(0.85)	(0.05)	(0.05)	(0.05)
Treated × Phase-out	−0.08*	−0.07*	0.45	−0.04	−0.14**	−0.16
	(0.04)	(0.04)	(1.02)	(0.10)	(0.06)	(0.10)
N	1758	1758	2295	2018	2227	2344
Firm FE		X				
Init. outcome	3.00	3.00	6.81	0.04	0.42	4.87
			Empirical strategy 2 – Matched training firms			
Treated × Roll-out	0.02	−0.01	−0.22	0.09	−0.15	−0.06*
	(0.04)	(0.04)	(1.02)	(0.09)	(0.11)	(0.04)
Treated × Post	0.03	0.01	−0.54	0.09	−0.12	−0.12*
	(0.05)	(0.04)	(1.07)	(0.07)	(0.08)	(0.07)
Treated × Phase-out	−0.03	−0.02	−0.85	0.08	−0.15	−0.08
	(0.05)	(0.05)	(1.41)	(0.11)	(0.11)	(0.13)
N	908	908	1233	1082	1190	1248
Firm FE		X				
Init. outcome	3.00	3.00	6.51	0.03	0.43	4.82

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

Table B2.2: DiD Results - Worker substitution effects continued

	Trainee retention rate	# VT separations	# VT hires	# low-educ. VT hires	# highly educ. VT hires
	(1)	(2)	(3)	(4)	(5)
	Empirical strategy 1 – All training firms				
Treated × Post	-0.10^{**}	4.19	0.27	0.58	-0.31
	(0.04)	(3.67)	(2.96)	(2.75)	(0.33)
	[-0.16;-0.03]	[-7.82;23.07]	[-8.00;9.66]	[-7.20;9.46]	[-1.11;0.09]
N	2259	2281	2295	2295	2295
Init. outcome	0.42	21.06	17.44	15.47	1.96
	Empirical strategy 2 – Matched training firms				
Treated × Post	-0.10^{**}	3.00	-0.91	-0.58	-0.33
	(0.05)	(4.04)	(3.07)	(2.80)	(0.38)
	[-0.16;-0.03]	[-8.16;16.83]	[-6.90;2.36]	[-5.91;2.30]	[-1.03;0.16]
N	1210	1224	1233	1233	1233
Init. outcome	0.42	18.93	17.79	15.77	2.03

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. Column 1: The trainee retention rate is equal to the share of trainees (independent of school education) which are offered a working contract after training graduation. VT: completed vocational training. For the main table, see Table 5.

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

	Investments per worker		Intensive vs. extensive margin				
	per init. # of workers	per current # of workers	Any inv. (0/1)	Log(Inv.)	Large inv. (1/0)	Combined	Log(K)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Empirical strategy 1 – All training firms						
Treated × Roll-out	−1.25	0.07	0.01	−0.29	0.00	−0.20	0.01
	(1.61)	(2.44)	(0.04)	(0.21)	(0.06)	(0.26)	(0.04)
Treated × Post	−3.36*	−1.00	−0.05	−0.33**	−0.07	−0.64***	−0.08
	(1.85)	(2.49)	(0.03)	(0.16)	(0.04)	(0.24)	(0.05)
Treated × Phase-out	−3.18	−3.21	−0.07	−0.12	−0.02	−0.61*	−0.05
	(2.72)	(3.65)	(0.05)	(0.21)	(0.07)	(0.35)	(0.08)
% of init. outcome	-19%	-6%	-6%		-12%	-10%	
N	2344	2344	2344	2069	2069	2344	2271
Init. outcome	17.68	17.67	0.89	7.46	0.56	6.63	10.31
	Empirical strategy 2 – Matched training firms						
Treated × Roll-out	−1.17	1.34	0.01	−0.33	0.01	−0.25	0.15*
	(1.90)	(2.68)	(0.05)	(0.24)	(0.08)	(0.33)	(0.09)
Treated × Post	−6.55***	−4.20	−0.03	−0.38*	−0.06	−0.48	−0.07
	(2.36)	(3.48)	(0.05)	(0.22)	(0.07)	(0.37)	(0.15)
Treated × Phase-out	−5.14*	−7.44	0.00	−0.26	−0.09	−0.21	0.02
	(3.01)	(6.17)	(0.07)	(0.29)	(0.08)	(0.55)	(0.17)
% of init. outcome	-37%	-23%	-3%		-11%	-5%	
N	1248	1248	1248	1102	1102	1248	1191
Init. outcome	17.93	17.93	0.88	7.41	0.57	6.55	10.26

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. For the main table, see Table 6.

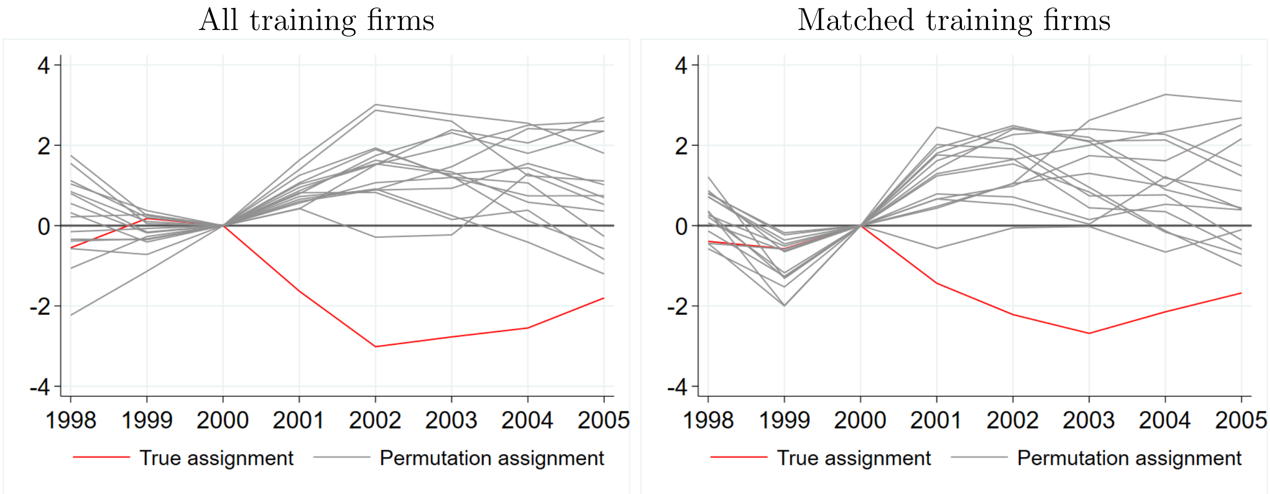
Table B2.4: DiD Results – Investment effects continued

	Investments per worker			Large investments		
	Overall	Business serv. + Public admin.	Manufacturing	Uppest tercile	Uppest decile	Uppest industry-specific tercile
	(1)	(2)	(3)	(4)	(5)	(6)
			Empirical strategy 1 – All training firms			
Treated × Post	-3.36^{*}	-1.87	-3.63^{**}	-0.07	-0.02	-0.07^{*}
	(1.85)	(3.55)	(1.39)	(0.04)	(0.02)	(0.04)
	[-6.91;-0.06]	[-8.83;3.57]	[-5.52;-0.36]	[-0.09;-0.03]	[-0.04;0.01]	[.]
N	2344	1040	808	2069	2069	2069
Init. outcome	17.68	24.61	6.49	0.56	0.08	0.36
			Empirical strategy 2 – Matched training firms			
Treated × Post	-6.55^{***}	-10.41^{**}	-3.01	-0.06	-0.06^{*}	-0.13^{**}
	(2.36)	(4.09)	(1.90)	(0.07)	(0.03)	(0.05)
	[-10.23;-3.21]	[-15.98;-5.33]	[-6.75;0.37]	[-0.16;0.07]	[-0.09;-0.02]	[.]
N	1248	672	336	1067	1067	1067
Init. outcome	17.93	24.61	6.49	0.57	0.09	0.36

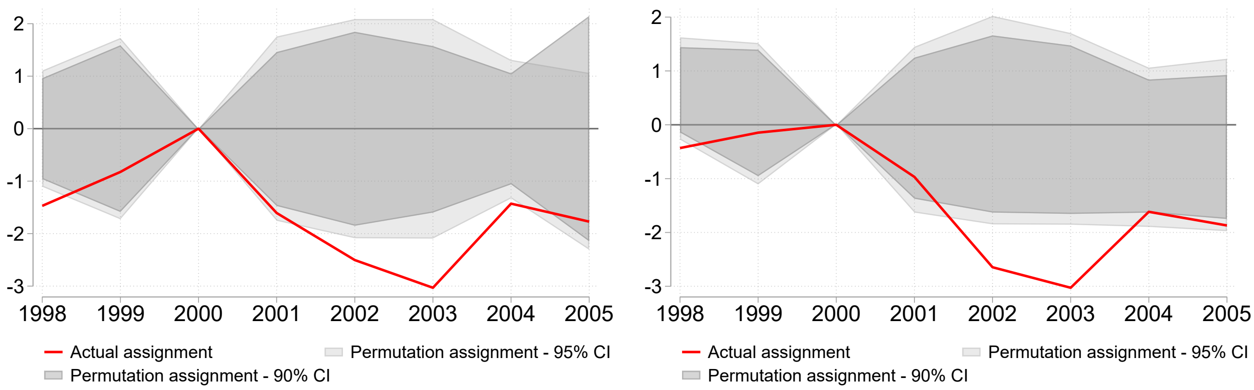
Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. Column 4: a commuter is defined as a person living and working in two distinct federal states. This variable is available from 1999 onward only. Column 5: Internal retraining is the sum of individual retraining incidences at the firm-year level. VT: completed vocational training. For the main table, see Table 6.

Figure B2.4: Permutation tests

A. Stock of highly educated trainees



B. Investments per worker in €1,000 – West Germany



Notes: T-statistic of the event study coefficients of the actual treatment assignment (red) and all potential permutation assignments (gray). *Panel B:* Permutation assignments within West Germany. For the main figure, see Figure 8.

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

	Technical status	Organizational change	Investment type (0/1): Production facilities	Investment type (0/1): ICT	Investment type (0/1): Real estate	Investment type (0/1): Transport *(Placebo)*
	(1)	(2)	(3)	(4)	(5)	(6)
			Empirical strategy 1 – All training firms			
Treated × Roll-out	−0.03	−0.10	−0.02	−0.06	−0.09*	−0.04
	(0.07)	(0.14)	(0.06)	(0.05)	(0.05)	(0.06)
Treated × Post	−0.18**	−0.37**	−0.07	−0.08*	−0.07	−0.01
	(0.09)	(0.16)	(0.05)	(0.04)	(0.04)	(0.05)
Treated × Phase-out		−0.27	0.00	−0.106	0.03	−0.02
		(0.17)	(0.07)	(0.06)	(0.06)	(0.06)
N	2341	1311	2344	2344	2344	2344
Init. outcome	3.97	1.35	0.70	0.79	0.58	0.33
			Empirical strategy 2 – Matched training firms			
Treated × Roll-out	0.06	−0.13	−0.06	−0.05	−0.06	0.00
	(0.09)	(0.22)	(0.07)	(0.06)	(0.06)	(0.07)
Treated × Post	−0.14	−0.66***	−0.09	0.00	0.00	−0.02
	(0.13)	(0.22)	(0.07)	(0.06)	(0.07)	(0.07)
Treated × Phase-out		−0.27	0.03	0.00	0.10	−0.04
		(0.24)	(0.09)	(0.09)	(0.10)	(0.07)
N	1245	702	1248	1248	1248	1248
Init. outcome	3.98	1.41	0.69	0.78	0.56	0.32

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. Technical status: Technical status of a firm's machinery on a scale from 1 ('completely out-of-date.') to 5 ('state-of-the-art'). Organizational change: On a scale from 0 to 4 as the sum of up to four organizational measures. ICT: Information & Communication Technologies. For the main table, see Table 8.

Table B2.6: DiD Results – Effects on firm-level technological change in non-training firms

	Technical status	Organizational change	Investment type (0/1) Production facilities	ICT	Real estate	Transport *(Placebo)*
	(1)	(2)	(3)	(4)	(5)	(6)
			Empirical strategy 1 – All training firms			
Treated × Post	−0.04	0.03	0.00	−0.04	0.01	0.01
	(0.04)	(0.06)	(0.02)	(0.02)	(0.02)	(0.02)
	[-0.13;0.05]	[-0.07;0.13]	[-0.05;0.04]	[-0.09;0.02]	[-0.03;0.05]	[-0.02;0.04]
N	8744	4737	8744	8744	8744	8744
Init. outcome	3.78	0.91	0.60	0.62	0.34	0.38
			Empirical strategy 2 – Matched training firms			
Treated × Post	−0.05	0.03	−0.01	−0.04	−0.02	−0.02
	(0.06)	(0.09)	(0.03)	(0.03)	(0.03)	(0.03)
	[-0.17;0.04]	[-0.15;0.35]	[-0.08;0.07]	[-0.10;0.01]	[-0.09;0.05]	[-0.07;0.01]
N	6112	3314	6112	6112	6112	6112
Init. outcome	3.78	0.91	0.60	0.62	0.34	0.38

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Post: 2002–2004. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Technical status: Technical status of a firm's machinery on a scale from 1 ('completely out-of-date.') to 5 ('state-of-the-art'). Organizational change: On a scale from 0 to 4 as the sum of up to four organizational measures. ICT: Information & Communication Technologies. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. 90% confidence bands based on cluster wild t-bootstraps in square brackets. Init. outcome: Average outcome of treated firms in 1998. . For the main table, see Table 8.

Table B2.7: DiD Results - Other outcomes

	Sales per worker (1)	Log employment (2)	Log wages (3)	Firm exit (4)
	Empirical strategy 1 – All training firms			
Treated × Post	-3.21	-0.12^{**}	0.00	0.00
	(26.08)	(0.05)	(0.01)	(0.01)
	[-84.55;39.60]	[-0.24;-0.05]	[-0.04;0.02]	[.]
N	1260	2344	2344	9141
Init. outcome	234.14	5.21	4.17	0.00
	Empirical strategy 2 – Matched training firms			
Treated × Post	-31.70	-0.11^{**}	-0.01	-0.01
	(37.09)	(0.06)	(0.01)	0.01)
	[-130.43;48.12]	[-0.23;-0.03]	[-0.05;0.03]	[.]
N	552	1248	1248	8104
Init. outcome	245.37	5.18	4.17	0.01

Notes: Reference group: Treated × Pre. Pre: 1998–2000. Roll–out: 2001. Post: 2002–2004. Phase-out: 2005. Controlling for period fixed effects, state fixed effects, Treated × 2000, and Treated × 2005. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Init. outcome: Average outcome of treated firms in 1998. Firm exit: Exit dummy taking the value of one for last year a firm exists (based on administrative data and not based on survey response) and zero otherwise. Each firm is included for all years the firm officially operated.

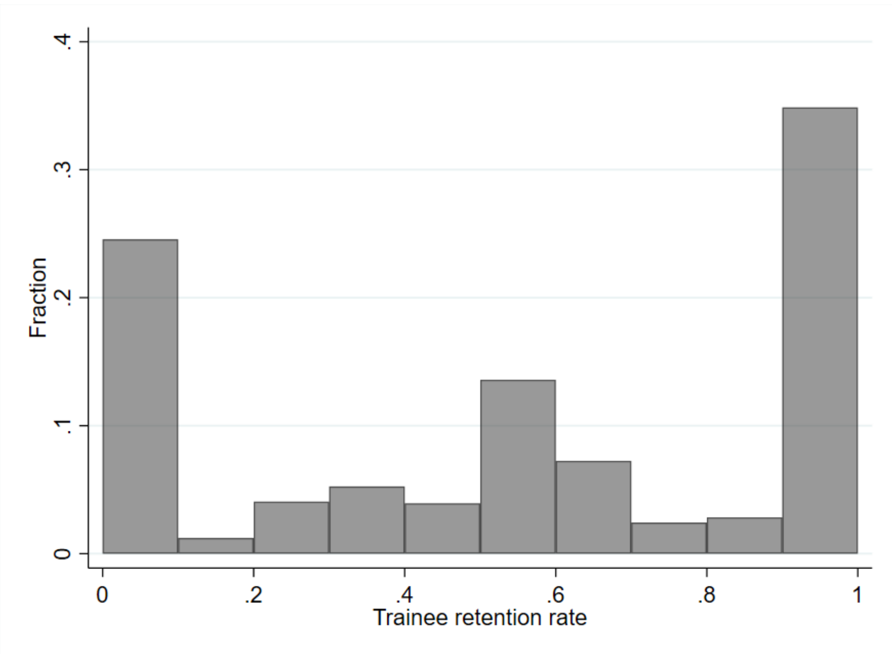
B.3 Additional results – Mechanism

Table B3.1: Heterogeneity by new skills

	(1)	(2)	(3)	(4)
New skills	-0.099	-0.100	-0.112	-0.113
	(0.135)	(0.129)	(0.130)	(0.124)
Industry	✓	✓	✓	✓
Exposure		✓		✓
Def. of new skills based on	1996-1999	1996-1999	2001-2003	2001-2003
Matched DiD N	78	78	78	78
Underlying total N	312	312	312	312

Notes: Outcome: Change in investments per worker in treated training firms compared to their matched control training firms between 2002 and 2000, see equation (2). New skills: Firm exposure to new skills measured as the 1998 share of workers in occupations with an updated curriculum between 1996–1999 (2001–2003). Between 1996-1999, 65 out of 232 occupations got updated; between 2001 and 2003, 47 occupations got updated. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For the corresponding main figure, see Figure 9.

Figure B3.1: Distribution of the trainee retention rate



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

Table B3.2: Heterogeneity by trainee retention rate

	(1)	(2)	(3)	(4)
Trainee retention rate	-6.14	-6.15	-13.76	-14.00
	(11.72)	(11.74)	(24.19)	(22.47)
Trainee retention rate^2			8.22	8.84
			(31.90)	(29.60)
Industry	✓	✓	✓	✓
Exposure		✓		✓
Matched DiD N	78	78	78	78
Underlying total N	312	312	312	312

Notes: Outcome: Change in investments per worker in treated training firms compared to their matched control training firms between 2002 and 2000, see equation (2). Firm trainee retention rate measured as the pre-reform share of trainees retained by the firm upon completion of the training. 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. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For the corresponding main figure, see Figure 9.

Table B3.3: Heterogeneity by firm employment growth

	(1)	(2)
Not shrinking	2.51	2.55
	(5.68)	(6.01)
Industry	✓	✓
Exposure		✓
Matched DiD N	78	78
Underlying total N	312	312

Notes: Outcome: Change in investments per worker in treated training firms compared to their matched control training firms between 2002 and 2000, see equation (2). Not shrinking: Dummy variable taking the value one if total firm employment does not decrease between 2000 and 2002. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. For the corresponding main figure, see Figure 9.

C Instrumental variable regression

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 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 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 trainee supply reduction, this complementary analysis identifies the causal effect of one additional trainee.

Main specification. 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 (C1). I instrumenting trainee employment as given in equation (C2):

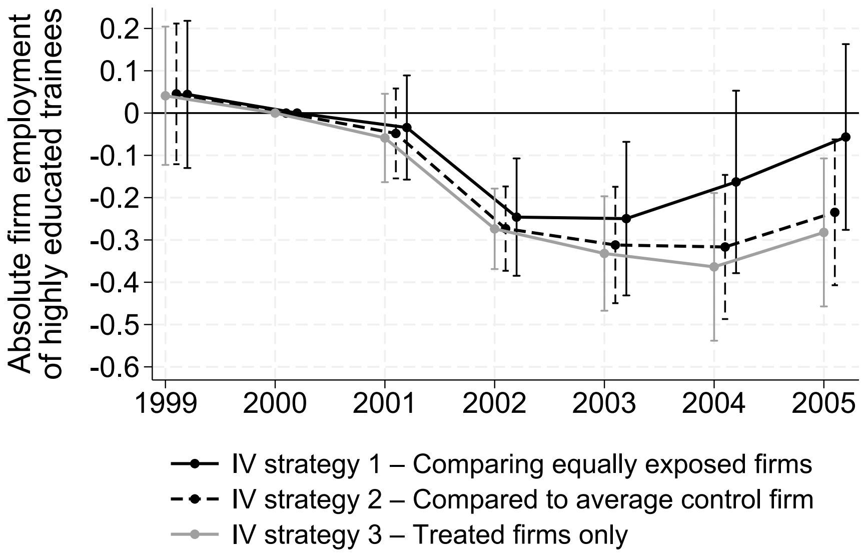
$$\begin{aligned} \text{Inv}_{jbt} &= N^{\text{Trainee}}_{jbt} + \psi_t + \pi_j + \epsilon_{jbt} && \text{(C1)} \\ N^{\text{Trainee}}_{jbt} &= \sum_t \gamma_t (N^{\text{Trainee}}_{j,1998} \times \text{Treated}_{b(j)} \times \text{Year}_t) \\ &\quad + \sum_t \zeta_t (N^{\text{Trainee}}_{j,1998} \times \text{Year}_t) + \psi_t + \pi_j + \epsilon_{jbt} && \text{(C2)} \end{aligned}$$

with j firms, b federal states, and t calendar years. Treated takes the value one if the firm is located in a state undergoing the education reform and zero otherwise. I predict trainee employment by firms' initial employment of highly educated trainees, i.e. firm exposure, $N^{\text{Trainee}}_{1998}$, corresponding to the *shares* in a shift-share instrument, times Treated × Year, corresponding to the reform-induced *shifts* in trainee supply across states and years. I control for nationwide time trends by firm exposure, $N^{\text{Trainee}}_{j,1998} \times \text{Year}_t$. Hence, the instrument exploits variation between two equally exposed firms located in a treated state and in 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 1998 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.[38] I run this regression among training firms, i.e. firms with initial exposures strictly greater than zero, to identify the effect of one additional trainee as opposed to the effect of having versus not having a trainee. Results are however very similar when running the regression among all firms, see Table C2, Panel B.

Figure C1 shows the coefficients of interest of the first stage, γ_t, (black solid line). Almost

[38] Results are robust to further restricting to the years 2000 onward, see Table C2, Panel A, but with smaller F-statistics.

Figure C1: IV results – First stage



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

perfectly in line with the event study estimates in Section 5, one additional trainee prior to the reform is associated with 0.25 trainees less in 2002 and 2003. This effect fades out in 2004 and 2005.

F-statistics and results of the second stage are displayed in Table C1, Panel A. Again, I report the effect on various specifications of the investment outcome, since the functional form of the relationship between investments and trainees is a priori unclear. With F-statistics of approximately 14, the instrument is relevant. Trainee shortages significantly decrease firm investments. In particular, one trainee less reduces investment by €550 per worker (column 1) or 4% (column 2). The probability to make a large investment decreases by 3 percentage points for each missing trainee (column 3) and log capital decreases by 2% (column 4). Again, confirming its relation to foregone technological change, measures of organizational change become less frequent (column 5), though the estimate is insignificant in this specification.

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 Table C1, Panel C. Convincingly, the results remain very similar.

Given that the instrument exploits reductions in trainee employment of on average -1.5 trainees per firm, the estimate delivers a credible treatment effect for similarly sized trainee reductions but are unlikely to be linearly scalable for substantially larger drops.

Alternative specifications. Trainee employment in control states might vary over time for reasons potentially related to investments. In this case, shifts among control firms are not exogenous. In an alternative specification, I therefore drop the term $N^{\text{Trainee}}_{j,1998} \times \text{Year}_t$ from the

Table C1: IV results – Second stage

	Inv. per worker (1)	Log(inv.) (2)	Large inv. (1/0) (3)	Log(K) (4)	Organizational change† (5)
	A. IV strategy 1 – Comparing equally exposed firms				
N^{Trainee}	0.55*	0.04*	0.03**	0.02***	0.05
	(0.31)	(0.02)	(0.01)	(0.01)	(0.04)
F–Stat	13.55	14.02	14.02	13.44	7.29
	B. IV strategy 2 – Compared to average control firm				
N^{Trainee}	0.60	0.07	0.04***	0.03***	0.16**
	(0.47)	(0.05)	(0.01)	(0.01)	(0.07)
F–Stat	14.25	14.68	14.68	14.40	9.07
N	2,051	1,798	1,798	1,991	1,066
	C. IV strategy 3 – Treated firms only				
N^{Trainee}	0.15	0.03	0.03***	0.02***	0.14**
	(0.44)	(0.05)	(0.01)	(0.01)	(0.06)
F–Stat	22.57	26.34	26.34	22.49	13.03
N	3,241	2,448	2,448	3,016	1,681

Notes: † – For data availability reasons, variable included for the years 2000, 2001, and 2004. F-Stat gives the robust Kleibergen-Paap Wald rk F statistic. Panel A and B: Training firms only. Panel C: All firms. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

first stage. This corresponds to comparing differently exposed treated firms to the average control firm each. Results are very similar to those from the main specification, see Figure C1, black dashed line, and Table C1, Panel B. The fact that the F-statistics do not decrease indicates that the omitted term does not have high explanatory power in the main analysis.

Variation in trainee employment in the main specification is both due to different exposures across firms and different shifts across states. To check whether more exposed firms are more affected by the shock and drop investments more, I next exploit variation due to different exposures only by restricting the set of firms to treated firms.[39] This alternative specification provides similar results as well, see Figure C1, gray solid line, and Table C1, Panel C. F-statistics markedly increase, consistent with the fact that the instrument is strong for firms in treated states, i.e. firms that actually experienced changes in trainee employment, but weak for firms in control states with no major changes in trainee employment.

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 con-

[39] I broaden the set of firms to include all firms instead of training firms only, to increase the number of observations. Repeating the same analysis among training firms relies on very few observations only but results remain comparable, see Table C2, Panel D.

Table C2: Robustness of IV results – Second stage

	Inv. per worker (1)	Log(inv.) (2)	Large inv. (1/0) (3)	Log(K) (4)	Organizational change† (5)
	A. IV strategy 1 – 2000–2005				
N^{Trainee}	0.51	0.06**	0.04**	0.02**	0.05
	(0.32)	(0.03)	(0.02)	(0.01)	(0.04)
F–Stat	8.84	7.87	7.87	8.63	7.29
N	1,758	1,520	1,520	1,707	1,066
	B. IV strategy 1 – All firms				
N^{Trainee}	0.43	0.03	0.03***	0.02***	0.06*
	(0.30)	(0.02)	(0.01)	(0.01)	(0.03)
F–Stat	15.48	16.62	16.62	15.19	8.42
N	9,702	7,705	7,705	9,184	5,053
	C. IV strategy 1 – Controlling for firm log employment				
N^{Trainee}	0.48	0.04	0.03**	0.02**	0.04
	(0.32)	(0.02)	(0.01)	(0.01)	(0.04)
F–Stat	12.29	11.99	11.99	12.17	6.82
N	2,051	1,798	1,798	1,991	1,066
	D. IV strategy 3 – Training firms only				
N^{Trainee}	−0.24	0.01	0.03**	0.01	0.13**
	(0.53)	(0.06)	(0.01)	(0.01)	(0.06)
F–Stat	24.07	28.31	28.31	24.32	11.02
N	567	484	484	560	295

Notes: † – For data availability reasons, variable included for the years 2000, 2001, and 2004. F-Stat gives the robust Kleibergen-Paap Wald rk F statistic. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

firms 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: each highly educated trainee employed less corresponds to approximately 4% investments less. This figure is lower than the one 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. firm investments decreasing in face of the trainee supply shock beyond the decrease related to foregone trainee employment. It might also hint at firm selection into trainee employment despite the shortage: If firms that would have invested in 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.

D Economic framework

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

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

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

For simplicity, assume that the intermediate outputs or tasks $y_{j\tau}$ are perfect substitutes.[40] As in Acemoglu (1998), technologies require skills, i.e. only workers trained for a specific technology, L_τ, can handle this technology. This production function zooms in on the relevant channel, namely the labor reinstatement channel of new technologies, i.e. the aspect that new technologies create new tasks performed by humans. It abstracts from other channels potentially occurring simultaneously, for example the automation channel, i.e. the aspect that new technologies automate tasks previously performed by humans. 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 $t = 1$ in order to maximize profits. In order to adopt the new technology τ, firms (re-)train a fraction Ψ_{τ_0} of workers of each initial productivity type. Training uniformly takes one period across technologies and workers.[41] Since workers within a cohort are homogeneous, firms always either retrain all or no worker of one entry cohort, $\Psi_{\tau_0} = \{0; 1\}$, such that worker cohorts and worker productivity types coincide. Wages W_τ are in proportion to, but below worker productivity due to firms' monopsony power, $W_\tau = \theta A_\tau$ with $\theta \in (0, 1)$. Benefits from technology-induced productivity

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

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

increases are hence not completely passed on to workers.[42] For now, assume that workers do not switch firms.

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

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

Firm maximization problem. Given the additive separability of intermediate outputs 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 $\tau_0 \in [0, ..., \mathcal{T} - 1]$. Additional profits are equal to the surplus in output minus the surplus in wages in period $t = 2$, henceforth "net output surplus", minus capital adjustment costs:

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

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

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

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

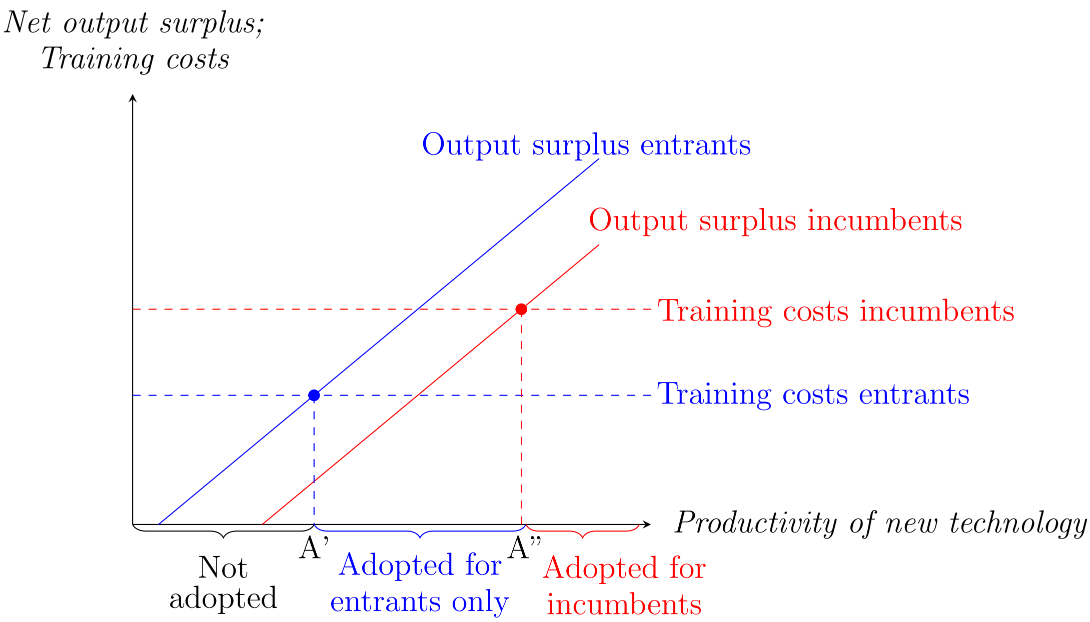
$$A_\tau \geq \left(1 + \frac{1}{1 - \theta}\right) A_{\tau_0} \tag{D5}$$

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

Assume there is a missing entry cohort in $t = 1$. In this period, firms invest in the new

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

Figure D1: Additional profits versus additional costs of technology adoption

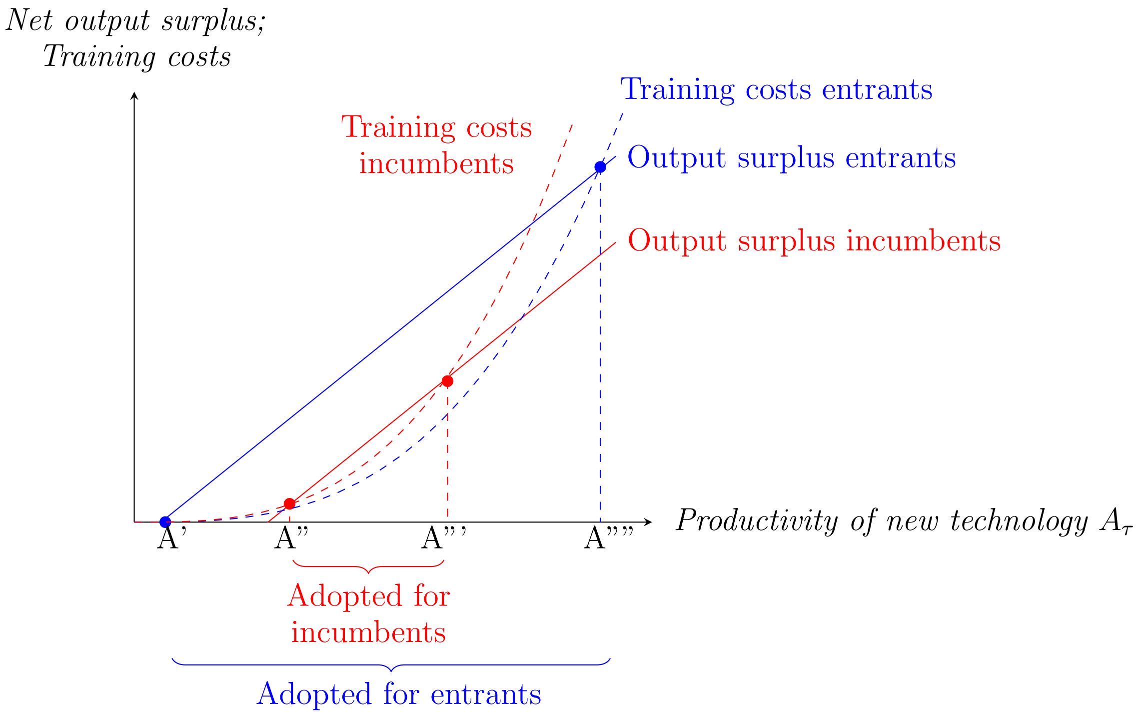


Notes: Profitability of training entrants versus incumbent workers.

technology τ if and only if the new technology is productive enough to make it profitable to retrain the least productive incumbent, i.e. if equation (D5) holds for $A_{\tau_0} = \min(A_{\text{incumbents}})$. For productivity levels of the new technology $A' \leq A_\tau < A''$, this implies a reduction in firms' technology adoption compared to the case without a missing entry cohort. Note that highly productive technologies above A'', that arrive very seldomly, are always adopted, also in times of a shortage of young labor market entrants.

Extension A – Increasing and convex capital adjustment costs. Equation (D2) implies constant capital adjustment costs for any productivity level of the new technology, namely one period of training. 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 new trade-off between additional profits and additional costs of production are shown in Figure D2. While the trade-off looks similar for small productivity levels of the new technology, under this assumption, there are new technologies in the upper tail of the productivity distribution that require prohibitively long training, exceeding the productivity gains. The productivity level above which adoption is prohibitively costly is reached earlier for incumbent workers because their training is more expensive. Hence, a new technology is adopted for entrants only below a certain productivity threshold A'' and above a certain productivity threshold A'''. In consequence, a lack of entrants not only hinders the adoption of technologies in the range between A' and A", as in the setting with constant training costs, but also of very productive technologies with productivities above A"'.

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



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

Extension B – Infinite time horizon and worker retention. I now allow firms and workers to live for an infinite time horizon and workers to switch firms.[43] For simplicity, assume there is no temporal discounting or capital depreciation. The expected total surplus of adopting a new technology τ is now given by the sum of all expected future net output surpluses minus one-time capital adjustment costs:

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

Workers can leave their firms at the end of each period. The probability of a worker to stay at a firm j, p_j, is exogenously given by a firm's monopsony power.[44] For each cohort, the net outplus surplus from technology adoption extincts as soon as this worker group is retrained in a new technology. The retraining probability $\phi(p_j)$ increases in the firms retention rate p_j, $\phi(p_j) > 0$. Hence, the expected net output surplus for each cohort is equal to the net output

[43] I abstract from worker retirement which would present an additional factor why training entrants is in expectation more profitable than retraining incumbents.

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

surplus of technology adoption in each (future) period multiplied by the probability of being at the firm and not being (re-)trained in this period. The expected total net output surplus is the sum across all cohort-specific net output surpluses:

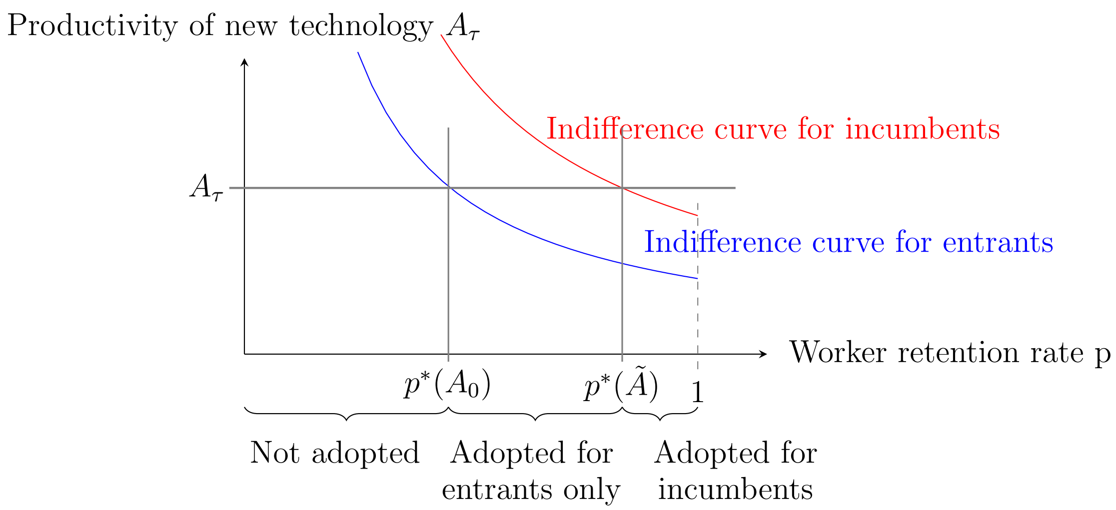
$$E[\Delta Y_{j\tau} - \Delta W_{j\tau}] = (1-\theta) \sum_{t+1}^{T} p_j^t (1-\phi(p_j))^t \sum_{\tau_0=0}^{\mathcal{T}-1} \Psi_{\tau_0} L_{jt\tau_0}(A_\tau - A_{\tau_0}) \tag{D7}$$

Technology adoption is more profitable the longer a firm benefits from trained workers, i.e. the higher p. Combining (D6) with (D2) and (D7), firms train a worker with initial productivity A_{τ_0} as long as the following condition holds:

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

For a fixed retention rate, the trade-off looks exactly the same as in Figure D1. Figure D3 visualizes the indifference curve along a firms' worker retention rate p, both for entrants with productivity A_0 (orange curve) and incumbents with productivity $\tilde{A}$ (red curve) for a given productivity level of a new technology at A_τ. This new technology is not adopted if the firms' retention rates is below the lower threshold $p^*(A_0)$. For retention rates above $p^*(A_0)$ but below the upper threshold $p^*(\tilde{A})$, the technology is adopted for entrants only. For retention rates above $p^*(\tilde{A})$, the technology is adopted for incumbents as well.

Figure D3: Indifference curves for Extension B



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

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