No Teens, No Tech: How Shortages of New Entrants Hinder Firm Technology Investments

Cäcilia Lipowski*

April 6, 2025

Abstract

This paper studies the role of young labor market entrants, particularly vocational trainees, in firm technology adoption. When new technologies require new skills, workers' ability to learn and adapt determines how well they complement these technologies, making new entrants particularly complementary due to their low opportunity costs and high expected returns from skill acquisition. Leveraging a large, temporary, and exogenous shock to trainee supply caused by an education reform in Germany in 2001, I provide causal evidence that a reduction in trainee supply decreases firm technology investments, indicating complementarity between new entrants and new technologies.

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

JEL: D22, D24, J21, J24, O33

---

*ifo Institute and LMU Munich, lipowski@ifo.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, Guido Friebel, Katja Görlitz, Maarten Goos, Simon Jäger, Morten Olsen, Harald Pfeifer, Pascual Restrepo, Johannes Schmieder, Fabian Waldinger, 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), CESifo Summer Institute on Re-skilling and Skill Shortages (Venice), Leading House Conference on the Economics of Educational and Vocational Training (Zurich), at seminars at University of Bonn, Frankfurt School of Finance and Management, ifo Munich, IWH Halle, IZA, LMU Munich, University of Mannheim, Rockwool Foundation Berlin, University of Strasbourg, Utrecht University and ZEW Mannheim, and at numerous other conferences and seminars. The project was mainly written at the ZEW Mannheim and 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

Whether technologies complement or substitute labor has been a central focus of research, with labor commonly classified by either skill levels or job tasks (e.g. Autor et al., 2003; Acemoglu & Autor, 2011). Following this reasoning, young workers, who tend to hold jobs with fewer high-skill, non-routine tasks, are less likely to be complements to new technologies. However, when new technologies demand skills that workers previously lacked, workers' ability to learn and adapt—rather than their initial skill level or task assignment—may become the primary determinant of their complementarity with new technologies. In that case, young labor market entrants play a key role in complementing new technologies: they have low opportunity costs and high expected returns to skill acquisition (e.g. MacDonald & Weisbach, 2004; Autor & Dorn, 2009; Cavounidis & Lang, 2020).

In this paper, I provide causal evidence of a complementary relationship between new entrants and firms' adoption of new technologies, with "new entrants" referring to young, not-yet-trained workers, and "new technologies" understood as potentially labor-saving technologies requiring new skills. I exploit a unique natural experiment in which an exogenous shock temporarily cuts off the supply of new entrants, while overall labor supply, population size, and labor demand remain stable. This setting allows me to isolate the role of new entrants, in contrast to studies of broader labor supply shocks—–such as migration or natural disasters (e.g. Lewis, 2011; Hornbeck & Naidu, 2014)—and avoids endogeneity concerns in firms' technology adoption that arise when comparing its impact on younger versus older workers (e.g. Aubert et al., 2006; Battisti et al., 2023; Barth et al., 2023; Aghion et al., 2024).

I begin with two stylized facts that suggest a complementarity between new entrants and new technology. First, in a representative firm survey, half of the firms offering vocational training state to do so because it substantially helps adapting to technological change and ensures the constant supply of new skills, while less than 20% negate it. Second, young workers are significantly more likely to work with new technologies than workers aged 30 and above, suggesting cohort effects in technology use. Technology use is highest among middle-skilled workers with vocational training, suggesting that this group plays a central role in technology *adoption*—–unlike the process of technology invention, which primarily involves high-skilled labor.

Drawing from these stylized facts, I sketch a simple economic framework that produces complementarity between young labor market entrants and new technologies through what I call the "new-skills" mechanism: Consider a task-based model where operating a new technology creates a new tasks requiring a new skill. Take for example office technologies used to manage orders, inventory, production schedules, and customer interactions. These technologies automate data entry, augment labor in decision-making, and simultaneously create new tasks requiring new skills related to system navigation or configuration of automated dashboards. In the model, firms assign new tasks to workers with the most profitable cost-benefit ratio for skill acquisition. Compared to incumbent workers, young labor market entrants face lower opportunity costs and greater productivity gains of learning new skills, since their initial productivity is low. Firms thus choose to complement technology adoption with new entrants, making technology use cohort-specific. When new entrants are scarce, retraining incumbent workers is often not profitable in which case forego the adoption of new technologies.

To identify the causal effect of new entrants on firm technology adoption—the main focus of this paper—I exploit a natural experiment arising from an education reform. From 2001 onward, two East German federal states (henceforth "treated states") increased the length of schooling required for the university entrance qualification by one year.[1] As a side effect, this reform lead to a one-time missing school graduation cohort from the upper school track in 2001, and subsequently a missing cohort of vocational trainees. There was no comparable drop in graduates or trainees in the other four East German states, henceforth "control states". In addition to being exogenous, this shock has a number of beneficial features: First, its temporary nature allows me to isolate the effect of trainee supply free from equilibrium adjustments—like labor shortages overall or wage increases—that a long-term decline in graduates would likely trigger. Second, the shock not only reduces the supply of trainees but cuts it off entirely – making the study high-powered to detect effects. Third, the reduction in trainee supply occurs without a concomitant demand shock because the total number of consumers remains unchanged.[2]

Although the missing school graduation cohort also affects the supply of university graduates several years later, identification is sharper for the negative supply shock of vocational trainees for two reasons: First, unlike university students, vocational trainees complete their programs after a fixed number of years, typically three. Second, vocational trainees tend to start their apprenticeships and first jobs close to their hometowns, unlike university students who often relocate for both their studies and first job.

In Germany, vocational training takes place in firms and vocational schools simultaneously. Trainees often remain at their training firm after completing the program. Workers with a vocational training degree make up the largest share of the workforce with two-thirds.[3] Trainees from the upper school track studied in this paper, i.e. those with 12 or 13 years of schooling, henceforth "highly educated trainees", make up 16% of all trainees (Federal Statistical Office, Genesis-Online, 2022a) while the majority have only 9 or 10 years of schooling, henceforth "low-educated trainees." Highly educated trainees often work in white-collar occupations such as media, retail, or financial services, which commonly require bachelor's or associate degrees in countries like the US. In general, vocationally trained workers are considered middle-skilled

[1] Among others, Büttner & Thomsen (2015); Morin (2015); Muehlemann et al. (2022); Marcus & Zambre (2019) and Dorner et al. (2024) 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] The composition of consumers changes, with some consumers being students instead of trainees in the reform year 2001. Since trainees earn very low wages, the difference in consumption between trainees and students should be limited.

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

professionals, rather than low-skilled laborers (Fitzenberger et al., 2024).

I compare investments and technology adoption of firms in treated East German states undergoing the temporary trainee shortage to those in control East German states in a difference-in-differences (DiD) event study design. I focus on training firms, defined as firms that used to employ trainees from the upper school track prior to the reform. The identification strategy rests on two main assumptions. First, absent the reform, firm outcomes in treated and control states would have followed parallel trends—a pattern supported by pre-treatment data. To rule out that concomitant industry-specific shocks drive the results, I match treated firms to comparable control firms within industries. Second, I assume that firms in control states are not affected by the reform. This assumption is likely to hold since vocational trainees are highly immobile (Muehlemann et al., 2022).[4] Consistent with this, I find no evidence of increased cross-state commuting after the reform.

Using a large and representative firm panel survey (IAB[5] Establishment Panel) linked with social security records (LIAB), I directly observe firm's trainee employment, investments in tangible assets plus ICT (information and communication technologies), and technical status of machinery. While the data lacks information about the specific technologies adopted, it comes with the advantage of encompassing a broad spectrum of investments and technologies rather than concentrating solely on one such as robots or computers.

I provide three key empirical findings. First, the education reform causes a temporary trainee shortage. Consistent with firms losing one entire trainee cohort and vocational training usually lasting three years, firms' employment stock of highly educated trainees decreases for three years. Supply of workers with completed vocational training is likely to be reduced for even longer, but this fails to be empirically verified. There is no anticipation effect in the form of increased trainee hiring prior to the shock, as trainee supply almost exclusively consists of fixed-sized school graduation cohorts, making it highly inelastic. Training wages do not increase, likely due to wage rigidities and the temporary nature of the shock. Firms do not compensate for the absence of highly educated trainees by hiring more low-educated trainees or workers with completed vocational training. This suggests that already trained workers are not effective substitutes for new labor market entrants.

The second key finding is that the trainee shortage reduces firm investments. Investments decrease sharply in training firms in treated states compared to control states during the years of the trainee shortage and catch up with those of control firms once the shock is over. There is no adjustment in firms' investments in anticipation of the trainee shortage. This can be rationalized by firms being unaware of the impending shortage—indeed, I find no evidence that treated training firms expected a shortage of new entrants—or by opposing firm responses that cancel each other out: some firms reduce technology adoption prior to the shortage,

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

[5] Institute for Employment Research

knowing that they would require several trainee cohorts to be trained on while just a few trainee cohorts will arrive before the shortage hits, while some firms invest more in anticipation. The investment decrease when trainees are missing is large: investments drop by 8% at the median (20% at the mean), or €1,140 per worker (€2,740). This is equivalent to a reduction by 0.4% (0.9%) of the capital stock. The sizable estimate is plausible considering that the intertemporal elasticity of long-term investments is nearly infinite: firms face almost no cost from delaying investments, but gain substantially by waiting until trainees become available again. Also, complier firms, i.e. firms that reduce their trainee employment when trainees are scarce, tend to be large and heavily investing, amplifying the effect. Indeed, when employing an additional identification strategy, i.e. a Bartik-type instrument exploiting pre-reform exposure to the shock, I find that each missing trainee is associated with foregone investments of €730 per worker.

The fact that firms time their investments to coincide with the availability of trainees informs the central take-away of the paper: Trainees and investments are complements. However, investments are unlikely to be generally labor-complementing, since overall labor was not scarce.

Investments drop in manufacturing as well as in business service firms, suggesting that a broad range of technologies are complementary with new entrants. The effect is driven by firms typically retaining their trainees post training, i.e. firms that tend to strongly invest in their trainees' human capital rather than using them as inexpensive labor. This provides evidence for the hypothesized channel of the complementarity: it is the trainees' role in acquiring new skills—rather than their role as low-cost labor—that drives the complementarity.

I confirm the link between the investment decline and the absence of trainees in two ways. First, comparing non-training firms across treated and control states, I do not find a comparable reduction in investments. Second, based on the auxiliary identification strategy exploiting firms' exposure to the shock, I show that firms which are more affected by the negative trainee supply shock decrease investments to a greater extent than less affected firms.

Third, the investment decline is linked to the reduced adoption of new technologies: the technical status of machinery depreciates in treated training firms compared to control training firms. However, firms undergoing the shortage manage to catch up with control firms after several years.

This paper contributes to three literatures. The most closely related strand of literature studies how technology invention and adoption respond to factor prices determined by their relative abundance (e.g. Zeira, 1998; Acemoglu, 1998, 2002). For example, an increased supply of skilled labor intensifies the adoption of skill-complementing technologies (Beaudry et al., 2010; Carneiro et al., 2022), while a decrease in the supply of (low-)skilled labor increases wages and incentivizes firms to patent and adopt labor-saving technologies (Lewis, 2011; Hornbeck & Naidu, 2014; Clemens et al., 2018; Dechezleprêtre et al., 2019; Danzer et al., 2024; Andersson et al., 2022; San, 2023, vice versa for an increase). The temporary shock studied in this paper leaves wage levels and long-term overall labor supply unaffected, allowing me to shut down the labor-saving channel, and to isolate the effect of the new-skills mechanism. This paper contributes to this literature in two dimensions. First, it highlights the role of training-related capital adjustment costs, which produce complementarity between new technologies and workers with a comparative advantage in skill acquisition (the "new-skills" mechanism). Second, it focuses on young labor market entrants—a decisive but previously overlooked worker group—and finds that even a temporary decrease in their supply causes investment responses.

Second, I contribute to the fragmented literature on new technologies, new tasks, vintage-specific human capital, and training, which points out that technologies create new tasks (Autor et al., 2024) requiring vintage-specific skills (Chari & Hopenhayn, 1991; Lipowski et al., 2024), and therefore worker training (Bartel & Sicherman, 1998; Bresnahan et al., 2002; Battisti et al., 2023), while simultaneously rendering skills of incumbents obsolete (Deming & Noray, 2020). In macro settings, vintage-specific tech skills have been put forward as the reason why technological change takes place through the entry of young workers, rather than through upskilling incumbent workers (MacDonald & Weisbach, 2004; Cavounidis & Lang, 2020; Adão et al., 2024), and why population aging above a certain threshold reduces technology adoption (Angelini, 2023). By finding evidence for technology use to be cohort-specific, this paper provides strong micro support for vintage effects which make firms forego technology adoption when new entrants are scarce.

In most general terms, this paper contributes to the extensive literature on the relationship between labor and technology. Most studies in this field examine whether workers and technology act as substitutes or complements by analyzing the labor demand effects of technological change. A subset of this literature focuses on older workers, finding that they are more likely to retire early, and receive no rents when firms invent or adopt new technologies (e.g. Aubert et al., 2006; Ahituv & Zeira, 2011; Battisti et al., 2023; Aghion et al., 2024). This paper identifies the relationship between new entrants and technology adoption in the reverse way, leveraging a clean identification strategy free from endogeneity concerns in technology adoption. In doing so, it provides direct evidence for firms timing their technology investments to coincide with the availability of new entrants, indicating complementary between the two.

The remainder of the paper is structured as follows. The next section presents two stylized facts motivating the following stylized economic framework which highlights the new-skills mechanism. Section 3 provides an overview of the German vocational training system and the education reform. Section 4 describes the data. I present the DiD event study approach in Section 5, followed by the empirical results regarding the reform's impact on trainee employment (Section 6) and firm technology investments (Section 7). Section 8 concludes.

2 Young labor market entrants and new technologies

2.1 Stylized facts

I provide two stylized facts that will inform the simple economic model. First, firms report offering vocational training to adapt to technological change and to ensure a steady supply of new skills. Second, young workers use new technologies more than older workers, suggesting cohort effects in technology usage.

Firm statements. The first stylized facts relies on stated preferences, suggesting that firms themselves view trainees and new technologies are complementary, potentially related to the relevance of new skills. I use the representative BIBB-Cost-Benefit firm survey, which asks firms engaged in vocational training about their reasons for offering it. Among all East German training firms surveyed in 2000, approximately half of the firms report using vocational training to substantially improve adaptability to technological change and to enhance the innovative capabilities, while only below 20% report not to do so. Similar shares are observed for firms using vocational training to ensure a constant supply of new skills and knowledge, see Table 1, suggesting the relevance of new skills as potential channel why trainees are important to adapt to technological change. For more details, see Appendix F.

Table 1: Use of vocational training according to firm survey

	Applies	Does not apply
Substantially improves adaptability to technological change	46%	19%
Enhances innovative capabilities	51%	18%
Ensures constant supply of new skills and knowledge	51%	16%

Notes: Based on the BIBB-Cost-Benefit Survey 2000. Firms in East Germany only. Responses range on a scale from 1 ("Does not apply at all") to 5 ("Fully applies"). Applies: categories 4+5. Does not apply: Categories 1+2. Using representative survey weights. N=521.

Technology use by age groups. The second stylized fact is that young workers use new technologies more than older workers, suggesting that cohort effects in technology adoption. I use a large, representative employee survey in Germany in 1999, 2006 and 2012 that questions respondents in Germany about their main working tools (IAB/BIBB/BAuA Qualification and Career Survey). In particular, I regress a binary variable on the usage of new technologies—computers and computer-controlled machines—on an age dummy, controlling for industry, occupation, year, education, and gender. Workers still in vocational training are unfortunately not included in the survey. Table 2 shows the results. I find that workers below the age of 30 are 4.2 percentage points (12%) more likely to mainly work with new technologies than workers aged 30 and above. This is a strong finding given that these workers are employed in the same industry and detailed occupation and have the same educational degree. For more information, see Appendix F. This finding is driven by workers with vocational training, see column 3, who have the highest overall usage level of new technologies, and who are 5.0 pp (13%) more likely to work with new technologies if they are below 30. This highlights the importance of studying middle-skilled professionals in technology *adoption*, since they are the ones actually *using* new technologies, in contrast to tertiary educated workers being relevant in technology *invention*. Similar results can be found for other European countries based on computer use at work reported in the European Working Conditions Survey, see Appendix F.

Table 2: Use of technology by age

	Overall (1)	No education (2)	Completed vocational training (3)	Tertiary educated (4)
$<$ 30 years	4.17***	3.42	4.96***	2.50
	(0.70)	(2.13)	(0.83)	(1.58)
Mean dep. variables	34.15	27.58	37.63	24.43
N	51,976	4,083	37,309	10,584

Notes: Outcome: Use of computers and computer-controlled machines (0/100). Based on the BIBB-BAuA Qualification and Career Survey 1999, 2006 and 2012. All regressions control for dummies for survey wave, gender, education level, occupations (353), industries (17). Heteroscedasticity-robust standard errors. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

2.2 Stylized economic framework

Guided by these stylized facts, I now present a simple economic model that rationalizes the complementarity between new entrants and firm investments in new technologies. The model builds on the tasks framework à la Acemoglu & Autor (2011), but assumes that each new technology introduces a new tasks requiring a new skill. Firms incur capital adjustment costs of worker training in this new skill, and assign the new task to workers with a comparative advantage in skill acquisition, making technology adoption endogenous to the supply of those workers.

Baseline setting. Firms operate and employees work in overlapping generations for two periods $t = 1, 2$, such that in each period, firms have access to two types of workers: entrants—untrained workers with a baseline productivity A_0—, and incumbents. Worker cohorts are unit-sized cohort and homogeneous. At the beginning of each period, a new technology vintage ν with productivity A_ν creating a task y_ν becomes exogenously available. Tasks are perfectly substitutable for simplicity.[6] Compared to the previous technology vintage $\nu - 1$, the new technology increases worker productivity by $\Delta A_\nu = A_\nu - A_{\nu-1}$, which follows a Poisson

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

distribution with a rate of 1, $\Delta A_\nu \sim Pois(1)$. Hence, technological progress is always positive, but rarely large. Only workers trained for a specific technology vintage ν—denoted L_ν—can handle this vintage.

In each period, a firm produces one final good Y under the following production function:[7]

$$Y = \sum_{\nu=0}^{\tau+1} y_\nu = \sum_{\nu=0}^{\tau+1} A_\nu L_\nu \tag{1}$$

with $\tau + 1$ the latest technology vintage currently available.

The price for the final product is fixed to one. Workers are identical in all aspects other than their cohort (skill and education level, occupation etc.) such that worker cohorts and worker productivity/technology types coincide. Wages w_ν are in proportion to, but below worker productivity, $w_\nu = (1-\theta)A_\nu$—with the wage wedge $\theta \in [0, 1)$ assumed to be exogenous—such that benefits from technology-induced productivity increases are not completely passed on to workers.[8] Let us assume for now that workers do not switch firms.

Firm maximization problem. Given perfect substitutability between tasks, firms maximize total profits π by independently deciding for each cohort ν whether to assign them the new task with productivity $A_{\tau+1}$ and provide the necessary training ($\mathbf{1}_\nu = 1$) or not ($\mathbf{1}_\nu = 0$). This produces cohort effects in technology adoption as the second stylized fact establishes. That is, for each cohort, the firm compares staying with the current technology vintage ν, which yields net output $Y_\nu - w_\nu L_\nu$ with adopting the new technology vintage $\tau + 1$, which yields $Y_{\tau+1} - w_{\tau+1}L_\nu - C_\nu$. Adoption costs C are borne by the firm and consist of foregone output during training, with training taking one period for all technologies and workers, $C = A_\nu L_\nu$. For each cohort, the firm solves:

$$\max_{\mathbf{1}_\nu \in \{0,1\}} \left[(Y_\nu - w_\nu L_\nu)(1 - \mathbf{1}_\nu) + (Y_{\tau+1} - w_{\tau+1}L_\nu - C_\nu) \cdot \mathbf{1}_\nu\right] \tag{2}$$

For each cohort, the firm adopts the new vintage ($\mathbf{1}_\nu = 1$) if and only if:

$$Y_{\tau+1} - w_{\tau+1}L_\nu - C_\nu > Y_\nu - w_\nu L_\nu \tag{3}$$

which can be rewritten as

$$\frac{A_{\tau+1}}{A_\nu} > 1 + \frac{1}{(1-\theta)} = \kappa \tag{4}$$

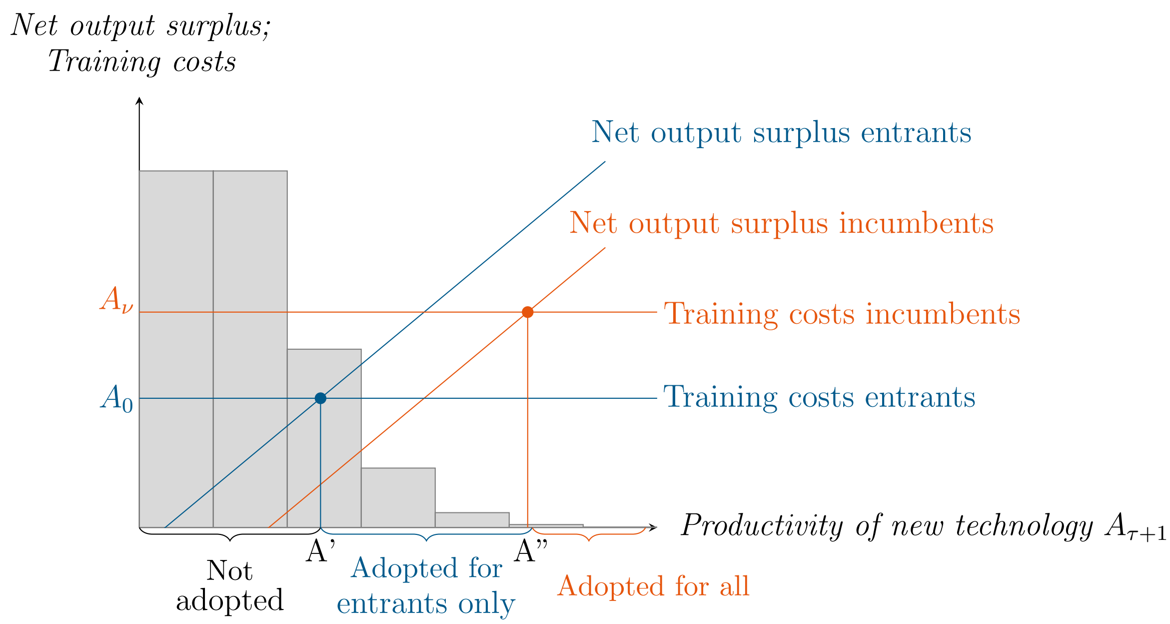
[7]This production function zooms in on the labor reinstatement channel of new technologies, i.e. new technologies creating new tasks performed by humans, while abstracting from automation or factor-augmenting forms which may happen simultaneously but do not constitute the central aspect of this paper.

[8]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).

Hence, a firm assigns the new technology-using task to a worker cohort if the productivity increase is above a certain threshold $\kappa = 1 + \frac{1}{(1-\theta)}$ with $\kappa \in [2, \infty)$ depending on the degree of labor market imperfections. The two factors determining the profitability of adoption are the wage wedge (the larger the share of workers' marginal product a firm can keep, the more likely the firm will adopt the new technology), and worker cohorts' initial productivity: The higher a worker cohorts' initial productivity, the lower the net output surplus, and the higher the training costs. Firms will therefore more often assign new technology tasks to entrants, making entrants essential for firms adaptability to technological change and the supply of new skills, as the first stylized fact establishes.

Figure 1 visualizes this trade-off. New technologies below the productivity threshold $A' = \kappa A_0$, i.e. technologies that are not at least twice as productive as entrants, are not profitable to adopt, not even for entrants, even though these rather unproductive technologies arrive frequently. New technologies above A' but below $A'' = \kappa A_\nu$, i.e. technologies that are not at least twice as productive as incumbent workers, are assigned to entrants only. New technologies above the threshold A'' are assigned to both incumbents and entrants. However, these highly productive technologies only arrive rarely.[9]

Figure 1: Firms' costs and benefits of technology adoption



Notes: Profitability of assigning new technology-using task to entrants versus incumbent workers. The histogram shows the productivity distribution of the new technology vintage $A_{\tau+1}$.

Missing entrant cohort. Assume that an entrant cohort is missing in $t = 1$. Firms invest in the new technology vintage $\tau + 1$ if and only if the new vintage is productive enough to make retraining incumbents profitable—that is, if the vintage is at least twice as productive

[9] In the case of increasing and convex capital adjustment costs, i.e. training costs, as the literature typically assumes, highly productive technologies are also only adopted by entrants because training costs are prohibitively large for incumbents to pay off, see Appendix E.

as incumbent workers. For productivity levels of the new technology $A' \leq A_{\tau+1} < A''$, this reduces firms' technology adoption compared to the case without a missing entrant cohort.

The reasoning can be extended to a dynamic setting in which firms operate for more than two periods. If an entrant cohort is missing in this setting, the trade-off firms face is between adopting the new vintage immediately or postponing adoption until the next period, when entrants become available.

Worker retention. We now relax the assumption that workers always stay at their firms by introducing the parameter $\rho \in [0,1]$, denoting the probability that a worker remains with their firm in the next period. Equivalent to above, for each cohort, the firm adopts the new vintage ($\mathbf{1}_\nu = 1$) if and only if:

$$\frac{A_{\tau+1}}{A_\nu} > 1 + \frac{1}{\rho(1-\theta)} \tag{5}$$

The higher the retention probability, the more likely a firm is to adopt a new technology, and, hence, the greater the potential for a missing entrant cohort to reduce technology adoption (see Figure E2).

Alternative mechanisms. There are at least two alternative explanations for the complementarity between young labor market entrants and technology adoption other than their low opportunity costs and high returns to skill acquisition. Both of them rest on workers' age. First, standard human capital theory suggests longer expected payoffs when investing in younger workers (the "horizon" channel in Cavounidis & Lang, 2020). Second, younger workers may generally possess more up-to-date tech skills.

Both channels are likely to be relevant. However, they cannot explain cohort effects in technology use because worker age only differs marginally across cohorts. Only opportunity costs and expected training payoffs dicontinuously vary across cohorts, as noted in Cavounidis & Lang (2020). For example, the Cost-Benefit Surveys of Vocational Training show that firm revenues from skilled labor activities of second-year trainees (third-year trainees) are 134% (254%) higher than for first-year trainees (Schönfeld et al., 2016, Table 18).

3 The German vocational training system and the education reform

Below, I describe the German vocational training system and the education reform.

3.1 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). Vocational trainees are regarded as future middle-skilled professionals and work in occupations that typically require bachelor's or associate's degrees in other countries, such as the US.

Adolescents usually start vocational training after graduating from one of the following three high-school tracks: the basic track (*Hauptschule*, 9 years of schooling) which qualifies students for vocational training in blue-collar occupations; the intermediate track (*Realschule*, 10 years) which prepares students for any vocational training, including training in white-collar occupations (media, financial services, or retail occupations, in addition to manufacturing and technical occupations); or the upper-track (*Gymnasium*, 12 or 13 years) which is required for university studies. Approximately a third of the upper-track school graduates choose to undergo vocational training,[10] such that in 2000, 16% of trainees had a university entrance qualification (*Abitur;* Federal Institute for Vocational Education and Training, 2009). Trainees rarely move or commute to their workplace: only 2% of vocational trainees move across states for their vocational training (SOEP, own calculations). Based on the data used in the subsequent analyses, the share of trainees commuting across states is similarly low at approximately 5%. After completing the vocational training, which typically lasts three years, a large share of trainees remain at their training company.

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 (1-2 days per week). This paper exclusively focuses on the on-the-job training part. Trainees are hired by their training company, and receive a wage that is usually subject to collective bargaining agreements and low.[11]

Vocational training is comparable to on-the-job training in other countries with two notable exceptions: First, in addition to on-the-job training at the firm, trainees receive vocational schooling, which equips them with skills that may be external to the firm. Second, regularly updated training curricula ensure that the training content includes up-to-date technical skills (Lipowski et al., 2024).

[10] 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.

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

3.2 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,[12] 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. I therefore assign Saxony-Anhalt and Mecklenburg-Western Pomerania as treated states and the other four East German states as control states. The education reform was decided in May 1996 in Mecklenburg-Western Pomerania and in January 1998 in Saxony-Anhalt.[13] By lengthening the years of schooling, the reform increased the level of education. More importantly, because the last cohort completing 12 years graduated in 2000 and the first cohort completing 13 years graduated in 2002, the reform resulted in a one-time missing upper-track school graduation cohort in spring 2001. Figure 2, 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 control states.

The education reform was initiated by the Social Democratic Party (SPD), which entered the government in both treated states in 1994.[14] The SPD advocated for 13 years of schooling to promote equal opportunity, while its main opponent, the Christian Democratic Union (CDU), pushed for a 12-year system to enhance the international competitiveness of German school graduates and improve the efficiency of the education system. I rule out that the governance of the SPD, or related policy or socio-economic changes, confound the effect of the education reform by comparing several state metrics between treated and control states before and after the reform, as well as between state-periods governed by the Social Democrats and those not governed by the Social Democrats, in Appendix B. In Appendix B, I also discuss why concomitant investment tax programs are unlikely to have confounded the trainee supply shock.

How does the missing school graduation cohort translate into the labor market? Usually, two-thirds of the missing upper-track school graduates eventually opt for university studies, while one-third eventually start vocational training. The missing school graduates of spring 2001 are hence expected to result in a missing entry cohort of highly educated trainees in fall 2001, and to reduce the stock of highly educated trainees for three consecutive years, 2001–2003, given that vocational training typically lasts three years. At that time, males in Germany

[12] Results are robust to excluding Brandenburg from the set of control states.

[13] For more information on the education reforms, see Kühn et al. (2013) and Helbig & Nikolai (2015). Between 2007 and 2013, all German federal states adopted to the 12-year system, with Saxony-Anhalt making the change in 2007 and Mecklenburg-Western Pomerania in 2008. To avoid potential confounding effects from these changes, this study ends in 2006.

[14] Brandenburg has always been governed by the SPD since reunification which is why they changed from 12 to 13 years immediately after reunification.

had to do military service of 10 months when reaching the age of 18, partly postponing the missing entry and prolonging the reduction in the supply of trainees by one year, i.e. until 2004.

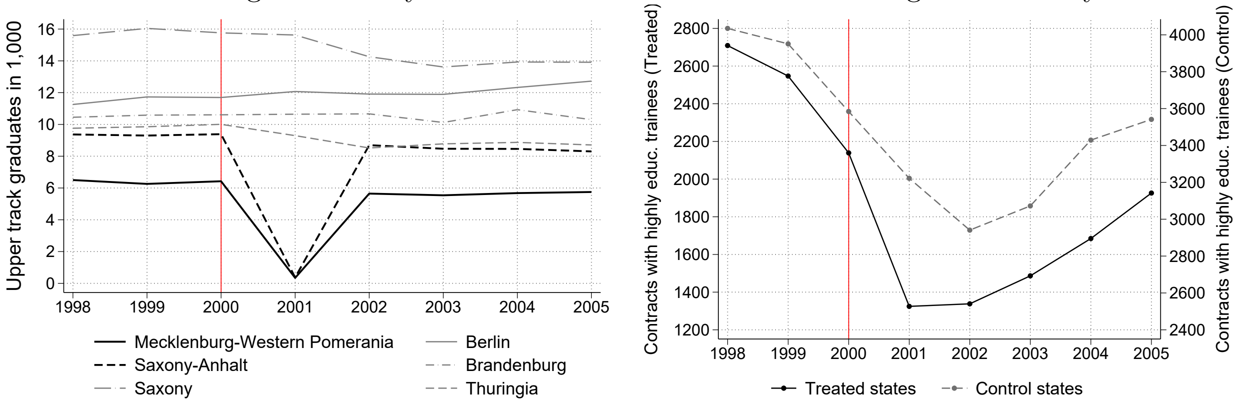
Official statistics confirm the decline in trainee supply: Figure 2, Panel B shows that training contracts with school graduates from the upper-track evolved in parallel in treated and control states between 1998 and 2000 but sharply fell in 2001 and 2002. While the decline is meaningful in magnitude, new training contracts with upper-track school graduates do not decrease by 100%, suggesting that school graduates from other cohorts partly compensate for the shock.

Dorner et al. (2024) analyzes how the education reform affects trainee employment. I instead use the temporary shock in trainee supply as the first stage, to study subsequent effects on investments. I focus on those upper-track school graduates who subsequently start vocational training instead of university students/graduates because vocational trainees postpone their labor market entry less and move or commute less across federal states, thus endorsing the credibility of the identification strategy. Note that the labor supply shock is unlikely to be confounded by a labor demand shock: the overall number of consumers remains unchanged, only the composition adjusts. Since trainees earn low wages, consumption patterns of trainees versus students is unlikely to have caused relevant demand changes.

Figure 2: The missing school graduation cohort

A. School graduates by state

B. New training contracts by state



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

Perhaps improving the credibility of the research design, both treated states are economically fairly different: Mecklenburg-Western Pomerania, located in the northeast of Germany along the Baltic Sea, is a predominantly rural and sparsely populated federal state with approximately 1.6 million inhabitants as of 2020. Its economy is defined by small and medium-sized enterprises engaged in agriculture, maritime industries, mechanical engineering, and tourism.

Saxony-Anhalt, situated in central Germany with a population of around 2.2 million in 2020, features a comparatively more urban environment. It is characterized by the chemical industry, mechanical engineering, and automotive supply. Both states, as well as control states, are characterized by excess trainee supply and high unemployment rates during this period, namely 17.8% in Mecklenburg-Western Pomerania and 20.2% in Saxony-Anhalt in 2000, see Figure C1, Panel B. Nonetheless, firms at the time report to experience severe skill shortages, in particular in ICT skills (e.g. Mitteldeutsche Zeitung, 2001).

4 Firm panel data

The analysis is based on the Linked-Employer-Employee-Data of the Institute for Employment Research (IAB), the LIAB-QM, which combines the annual representative IAB Establishment Panel survey with administrative employment information of all employees at surveyed firms.[15] The Establishment Panel has existed in West Germany since 1993 and in East Germany since 1996. The number of surveyed establishments has risen from 4,000 in 1993 to 16,700 in 2020. Importantly, the survey is conducted at the workplace level, enabling the distinction between treated and untreated establishments based on their location.[16] 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 the fall, such that trainees starting in fall each year appear in the data with a lag of one year. To account for this, I use employment figures from June of the following year to represent employment levels in the fall of the current year.

The data provide a reliable distinction between trainees and workers who have completed their training program, in addition to wages and employment status. Also, information on schooling allows me to distinguish "highly educated" from "low-educated" trainees, i.e. trainees with a university entrance qualification and those with a lower schooling degree, respectively. This is important since the education reform directly affects highly educated trainees only.[17]

I restrict the data in the following ways. First, I focus on the period 1997–2006. Second, I limit the data to firms in East Germany including Berlin, since the firms in East Germany are likely not comparable to firms in West Germany, especially for this period relatively shortly after reunification. However, results are robust to using the West German federal states as control states as well. Third, I exclude firms in the health/education/social service sectors, because vocational training in many related occupations is entirely school-based; as well as

[15] I use the cross-sectional model which comprises employment spells that encompass June 30 of each year. The LIAB longitudinal model is not suitable because it is only available for firms surveyed during 2009–2016.

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

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

firms in agriculture and arts/sports based on their extremely low share of highly educated trainees. Fourth, I drop very small firms, defined as those that with less than five employees in 2001, because they tend to exhibit volatile training investment behavior. Firm size is given by the number of full-time employed workers subject to social security contributions. Fifth, I only keep observations with non-missing investment values and firms that invest at least once throughout the entire time period. Sixth, I only keep firms that have existed in 1997. The firm panel is unbalanced: There is panel attrition in firms participating in the survey: 36% of firms observed in 1997 are still present in 2006. Estimates for later years must therefore be interpreted with caution.

The final sample comprises 2,246 distinct firms, of which 744 are treated (416 in Saxony-Anhalt and 328 in Mecklenburg-Western Pomerania) and 1,852 are untreated. Table 3 shows summary statistics of the final dataset. In sum, all firms cover almost 270,000 workers per year, amounting to approximately 3.6% of the East German workforce in a year.[18] I observe 15,010 trainees on average across years, of which 2,456 (16%) are highly educated. In 78% of the firm-by-year observations, no highly educated trainee is employed, and 62% of firms never employ a highly educated trainee over the entire time window 1997–2006.

Table 3: Summary statistics – Full data sample

	Mean	SD	Min	Max	Yearly sum
# workers	163.75	452.38	1	12,133	267,389
# trainees	9.19	52.10	0	3,181	15,010
# highly educated trainees	1.50	8.84	0	461	2,456
No highly educated trainee	.78	.41	0	1	1,276
No highly educated trainee ever	.62	.49	0	1	1,013

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

Training versus non-training firms. Since the reform affects highly educated trainee employment only, I focus on training firms, defined as firms with at least one highly educated trainee in fall 1995, 1996, or 1997 (based on the employment information in June 1996, June 1997, and June 1998). These years are chosen to minimize the chance that a firms' training status is affected by firms anticipation of the trainee shortage. This divides the sample into 571 training firms and 1,675 non-training firms. Training and non-training firms are fundamentally different. Training firms are more than four times as large in employment as non-training firms, operate more often in the business service and public administration sector, and less often in construction, retail/motor vehicles, and hospitality and other services.

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

Descriptive statistics based on the final matched sample of training firms (for a description of the matching procedure, see 5) are given in Table 4. Firms employ 294 workers on average, of which 23 (9%) are highly educated and have a completed vocational training degree or are still in vocational training.

Investments. Each year, firms are asked whether they invested in four investment types in the last year: (1) production facilities, plant and equipment, furniture and fixtures, (2) communication technology, electronic data processing, (3) real estate and buildings, and (4) means of transport, transportation systems. Unfortunately, the survey does not provide investment figures for each category separately. Instead, if a firm invested in at least one of these categories, the firm is surveyed on the total amount of annual capital investments. Accordingly, the investment volume is expected to contain investments in these four categories, while it is unlikely that intangible assets other than ICT are included. Table A1 provides a detailed description of the underlying survey questions.

I divide total investment by the number of workers to account for the fact that large firms tend to make large investments and to purge the distribution from the right-skewness caused by the right-skewness in firm employment size. To curtail the impact of extremely large investments, I winsorize investments per worker that exceed 4 times the interquartile range. Across firm-year observations, 92% have positive investments; the mean investment is €5.4million, or €14,500 per worker, while the median investment is €1.4million, or €7,600 per worker. The investment per worker distribution is hence still highly right-skewed and has a large standard deviation of €14,460 per worker.

The establishment panel lacks a direct measure of the capital stock. To fill this gap, I exploit information on total investments, the proportion of net investments, dummy variables representing the four investment types, and industry. I apply the modified perpetual inventory method developed by Müller (2008, 2017) explicitly for this dataset to impute the capital stock.[19] The average capital stock per worker is €360,750. Acknowledging the inherent inaccuracies in this method, I focus on investments while reporting results for the log capital stock only in order to assess the effect size.

Projected technological change. Unfortunately, information on investment types (production facilities, ICT, real estate, and transport) only distinguishes between "no" or "any" investments in this category. Since the subsequent analyses reveal that the intensive investment margin is much more affected than the extensive investment margin, these variables will not be used to study technology adoption. To determine whether investments incorporate new technologies, I therefore use one additional piece of information from the data: The technical

[19] I establish a starting value for the capital stock using investments in the first three observed years (1996, 1997 and 1998 at the earliest) and project the capital stock for subsequent years using investment information and sector-specific depreciation rates. Since the capital stock is therefore highly unreliable in the first three years, I assign missing to the capital stock in 1996–1998.

Table 4: Descriptive statistics – Training firms, matched sample

	p5	Median	Mean	p95	SD
# workers	21	199	294	828	353
# workers with+in vocational training	18	155	222	650	262
# highly educated workers with+in vocational training	1	11	23	87	30
% highly educated workers with+in vocational training	1	6	9	25	8
# highly educated trainees	0.00	1.00	4.13	17.00	6.84
Any investments (1/0)	0.00	1.00	0.92	1.00	0.27
Total investments in €1,000	0	1,425	5,410	19,388	11,174
Investments per worker in €1,000	0.00	7.60	14.54	37.50	14.46
Capital stock per worker in €1,000	2.28	143.25	306.75	781.55	313.97
Projected technical status of machinery	3.00	4.00	3.92	5.00	0.63

Notes: Averages across 1997–2000. Among training firms only. p5 – 5th percentile; p95 – 95th percentile; SD – Standard deviation.

status of a firm's plant and machinery.

Firms are asked to assess the overall technical status of their production equipment compared to other establishments in the same industry on a scale from 1 ("completely out-of-date") to 5 ("state-of-the-art"). The technical status is a stock variable, unlike investments, and current investments also affect a firms' future technical status. Given this time lag between investments and technical status, I construct a variable called "projected technical status" which attributes a firm's future technical status back to the year in which the investments influencing that status were made. To estimate how the technical status depends on current and past investments, I regress a firms' technical status on the sequence of current and past investments, controlling for year and firm fixed effects, see Table C1. The investments of the past three years mainly affect a firms' current technical status. I therefore compute the projected technical status as the weighted sum of the firm's technical status over the next three years, with weights equal to the estimated regression coefficients.[20] A projected technical status of 4.49 in 2001 may, for example, indicate that investments in 2001 are associated with a technical status of 4 in 2002 (weighted with an importance of 51%), and of 5 in 2003 (weighted with an importance of 32%) and 5 in 2004 (weighted with an importance of 17%). The average projected technical status is 3.92 with a standard deviation of 0.63.

5 Event study approach

The identification strategy exploits the quasi-random assignment of the education reform to federal states that entails exogenous variation in the supply of upper-track school graduates

[20] In particular, I compute Projected technical status$_t$ $= \beta_{t-1}/(\beta_{t-1}+\beta_{t-2}+\beta_{t-3})$Technical status$_{t-1}$ $+$ $\beta_{t-2}/(\beta_{t-1}+\beta_{t-2}+\beta_{t-3})$Technical status$_{t-2}$ $+ \beta_{t-3}/(\beta_{t-1}+\beta_{t-2}+\beta_{t-3})$Technical status$_{t-3}$ with β_{t-1} to β_{t-3} the regression coefficients from Table C1.

across states and years. I compare training firms in treated and control states before and after the reform in a difference-in-differences (DiD) event study design by estimating the following specification:

$$Y_{jt} = \sum_{t=1997, t \neq 2000}^{2006} \beta_t(\text{Treat}_j \times \text{Year}_t) + \psi_t + \phi_j + \epsilon_{jt} \tag{6}$$

where Y is one of several outcomes such as investments, j denotes the firm, and t the calendar year. I study firms up to 2006 because a different education reform affects trainee supply from 2007/2008 onward. Treat is a binary variable with Treat $= 1$ if the firm is located in a state undergoing the reform and zero otherwise. ψ_t captures calendar-year fixed effects. Firm fixed effects ϕ_j capture time-constant level differences between firms. The vector $\beta_t, t \geq 2001$ includes the coefficients of interest, namely the differential firm outcomes in treated states compared to control states following the reform in 2001 purged from the baseline difference between treated and control firms in 2000. The event study thus identifies the causal effect of a firm facing a state-wide negative trainee supply shock.[21] I follow Roth et al. (2023) for the most recent suggestions for DiD estimations. However, since treatment is not staggered, potential biases common to two-way fixed effects estimators in a staggered setting (e.g. Goodman-Bacon, 2021) are irrelevant here.

For brevity, I also estimate the equivalent DiD specification, assigning the years 1997–2000 as pre-period and the years 2001–2003 as post-period (dropping the years 2004–2006):

$$Y_{jt} = \delta(\text{Treat}_j \times \text{Post}_t) + \xi \text{Post}_t + \lambda_j + u_{jt} \tag{7}$$

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

I estimate equations (6) and (7) for training firms. The implicit assumption is that treated training firms would have wanted to continue training during the time of the shock in absence of the reform, and are therefore directly affected by the shock. In contrast, non-training firms are affected only via spill-over effects. I therefore rerun the regression for non-training firms as a falsification test and expect much smaller estimates.[22] Here, the implicit assumption is that treated non-training firms would not have wanted to train during the time of the shortage in the absence of the reform.

Matching. Treated training firms may differ from control training firms in aspects which expose them to different potentially confounding factors. To ensure that treated and control training firms are comparable, and therefore exposed to similar potential confounders, I match

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

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

firms based on their pre-treatment characteristics in two steps. First, I match firms within training and non-training status, and nine industry groups. Results are robust to matching on more detailed industry classes. By matching within industries, the estimated reform effects are devoid of distortions arising from industry-specific shocks. Second, I perform a Mahalanobis distance matching with replacement. This metric minimizes the standardized Euclidean distance of the matching variables between treated and control firms while taking into account the correlation between the matching variables. The matching variables include pre-treatment log overall employment, pre-treatment relative employment of highly educated trainees, both averaged over the years 1997–2000, and the increase in pre-treatment log employment between 1997 and 2000.[23] To avoid excessively limiting the sample size while ensuring good compatibility, I keep the three control firms with the smallest Mahalanobis distance for each treated firm and subsequently discard the worst 10% of all matches. Results are robust to both aspects.

Table 5 shows characteristics of treated training firms compared to control training firms for both the unmatched and matched sample. Prior to matching, treated firms are significantly smaller than control firms. The matching works well in eliminating differences in observable firm characteristics, both targeted and non-targeted ones.

Table 5: Matched training firms – Balancing table and descriptive statistics

	Unmatched		Matched	
	Δ Mean	SE	Δ Mean	SE
	A. Targeted variables			
Δ log employment	-0.09	0.05	-0.04	0.03
Log employment	-0.37	0.13**	-0.13	0.12
% highly educ. trainees in total empl.	-2.04	3.49	0.23	0.28
	B. Non-targeted variables			
# workers	-220	74***	-28	38
# highly educated trainees	-3.51	1.84	-0.29	0.72
Any investments (1/0)	-0.02	0.02	-0.00	0.02
Investments per worker in €1,000	1.13	1.19	0.53	1.47
Capital stock per worker in €1,000	48.77	28.78	35.81	34.42
Projected technical status	0.04	0.06	0.04	0.07
Number of firms	571		445	

Notes: Averages across 1997–2000. Δ log employment refers to the change in log employment between 1997 and 2000. SE: Standard error. Among training firms only. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

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

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

Assumption 1 – Parallel trends. First, I assume that firm outcomes in treated states would have evolved in parallel to those in control states in the absence of the reform . Check for parallel trends prior to the shock to evaluate the credibility of this assumption, as I do in Sections 6 and 7, reveals that this assumption is likely to hold.

In Appendix B, I show that key state metrics such as unemployment, population size, education expenditure, public debt and public investments do not change differently in treated compared to control states post 2000.Likewise, I argue in Appendix B that the concomitant investment tax programs studied in Lerche (2022) and Siegloch et al. (2025) are unlikely to confound the effect of the education reform. Moreover, one might be concerned that the introduction of the euro in 2002, the German Hartz reforms over 2003–2005, the bust of the dot-com bubble in 2000, or China's accession to the World Trade Organization in 2001/2002 might confound the reform effect. However, these shocks likely affected treated and control East German states similarly, especially within industries. In addition, it is unclear why any other shock would affect firm outcomes differently based on the share of highly educated trainees at a firm. Beyond these general arguments, I test whether states more strongly affected by the bust of the dot-com bubble, i.e. Berlin and Saxony, drive the results. Results are robust to their inclusion.

Assumption 2 – No anticipation. The second identifying assumption is that firms did not change their behavior prior to the reform. Since the reforms were decided in 1996 and 1998, firms had the opportunity to adjust their employment and investments prior to 2001. However, the event study estimates show little evidence of this.

Students may have also anticipated the reform. There was, however, very little scope for them to react: When the reform was decided, students of the missing graduation cohort were in grade 7 in Mecklenburg-Western Pomerania and grade 9 in Saxony-Anhalt. Since the choice of school track was due after grade 6 in East Germany, it was not impacted by the reform. Also, school graduates may delay or accelerate the start of their vocational training in response to the shock. This would bias the estimates toward zero.

Assumption 3 – No spill-overs/SUTVA. Third, I assume that control states are not affected by the reform, and treated states are not affected by the absence of the reform in control states. This assumption is violated if trainees move or commute across federal states. The data allows me to identify cross-state commuting. Trainees in the affected states rarely commute across states (2.7% in 1999 to 2001) compared to workers with a university degree (5.3%), and this share does not change in response to the reform, see Section 6. To investigate whether school graduates move for their apprenticeship, I turn to the Socio-Economic Panel (SOEP) which tracks individuals from childhood onward. The cross-state trainee mobility rate is extremely low at 2.2%. Further, there is no instance of a highly educated trainee relocating to one of the treated federal states in the post-reform years 2001, 2002 or 2003 in the data.

However, if trainees moved or commuted from control states to treated states in response to the reform, this would bias the estimates toward zero.

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

One other aspect of the reform is the increase in the skill level of highly educated trainees due to the increased years of schooling. Also, the share of upper-track graduates who start vocational training might have been affected. However, these two aspects changed permanently, such that effect dynamics will help distinguish between these permanent adjustments and the temporary trainee shortage. In addition, higher levels of education would, if any, likely induce more, not less, investments, and therefore provide a lower bound of the effect.

One potential concern is that the supply shock may be accompanied by changes in demand. This issue typically arises with labor supply shifts due to migration. However, I focus only on a postponement in the start of vocational training, which is unlikely to meaningfully affect consumer demand. First, because the overall population size remains constant. Second, since trainees earn low wages, students counterfactual demand–had they started vocational training instead of staying in school for an additional year–is unlikely to differ significantly.

Turning to firm demand, low trainee wages also prevent a meaningful decrease in the firm wage bill when trainee employment is reduced, making it unlikely to present a confounding channel.

Another concomitant aspect of the reform is the potential substitution of missing trainees with workers of a different observed or unobserved type.[24] However, I do not interpret such substitutions as a source of bias but as a mechanism via which the effect unfolds. Besides, I 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, for the investment outcome, $\hat{\beta}_t$ are small if trainees are primarily missing in firms that would not have invested in the absence of the shock, and $\hat{\beta}_t$ are large if trainees are primarily missing in firms that would have invested in the absence of the shock. In what follows, I directly study the characteristics of the compliers of the reform. Also, in order to identify the effect on investments independent of the realized distribution of trainees across firms, I propose a complementary identification strategy in Appendix D: I predict the distribution of trainees across firms based on a Bartik-style instrument of firms' pre-reform use

[24] 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.

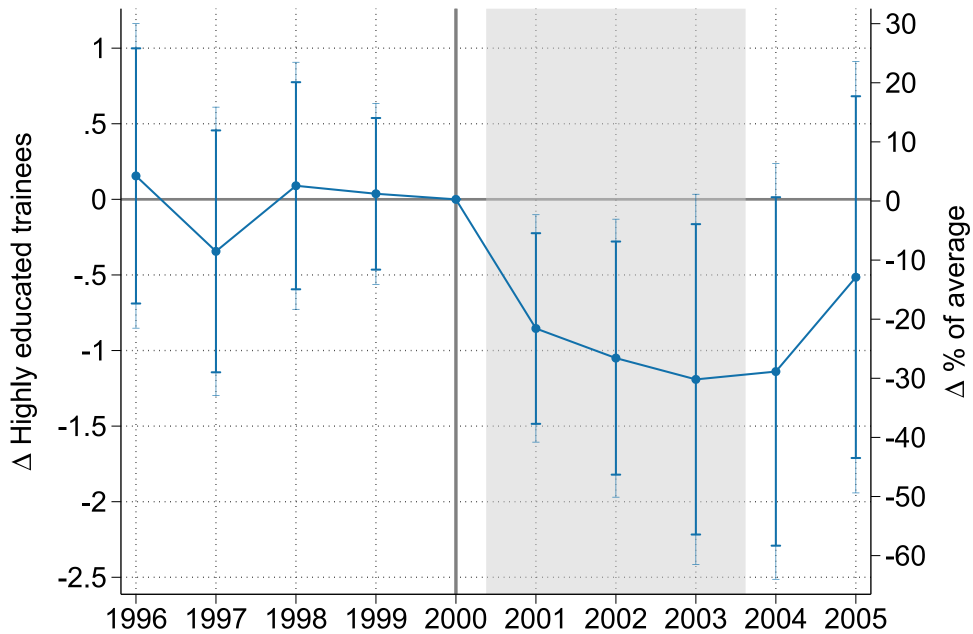
of trainees and the state-level shift in trainee employment induced by the reform. This allows me to identify a different causal parameter, namely the effect of employing one trainee fewer. This analysis is, however, more demanding and subject to further assumptions, which is why my preferred identification strategy is the DiD event study design.

Inference. Standard errors are commonly clustered at the level of treatment assignment to account for cluster-level shocks (e.g. Abadie et al., 2023). Here, this would result in a small number of clusters, i.e. federal states. For valid inference with a small number of clusters, I follow Roth et al. (2023), and assume that any state-specific shock is small compared to the idiosyncratic error terms at the firm level, potentially resulting in a small violation of parallel trends. This assumption is well justified in the data: For the two main outcomes, trainee employment and investments, the variance of the error term is approximately equal to the average variance of the error term within states but much smaller within firms, suggesting little to no within-state correlations but large within-firm correlations, see Table C2. I hence cluster standard errors at the firm level. Additionally, I perform permutation (Fisher randomization) tests, comparing the t-statistic of the treatment effect for the actual treatment assignment and all permuted treatment assignments across federal states.

6 Bite of the reform

Effect on trainee employment. Figure 3 displays the results of estimating the DiD event study model outlined in equation (6), i.e. the effect of the reform on the employment of highly educated trainees. Endorsing the identifying assumption of parallel trends, firms' highly educated trainee employment evolves in parallel in control and treated states in 1996–2000. This is expected, as trainee supply almost exclusively consists of fixed-sized school graduation cohorts, making it highly inelastic. During the years 2001–2004, employment of highly educated trainees is significantly lower in treated compared to control firms. Considering the typical training duration of three years plus one year delay for those going to military or social service, these are precisely the years the majority of the missing school graduates would have undergone vocational training. Consistent with the timeline of the shock, the employment gap shrinks in 2005 and becomes statistically not significantly different from zero. Absolute employment of highly educated trainees decreases by 1.01 on average in affected years. While this may sounds little, the effect corresponds to a reduction in the stock of trainees by a third, i.e. one out of three simultaneous trainee cohorts, for three years, or, put differently, to a decline by one entire new trainee cohort. Hence, the education reform not only reduces the supply of trainees—it eliminates it completely, making the setting high-powered to detect effects on firm adjustment strategies.

Figure 3: Effect on employment of highly educated trainees



Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (6). Left axis: Difference in the absolute employment stock of highly educated trainees per firm between treated and control firms. Right axis: Difference expressed in percent of the average pre-treatment absolute employment stock of highly educated trainees among treated firms. Standard errors clustered at the firm level. Among training firms only. N=3,609.

Wage and worker substitution effects. The detailed administrative labor market data allow me to study firms' adaptation strategies, such as changes in trainee wages, or the substitution of highly educated trainees with other workers. To investigate such effects, I employ the corresponding DiD specification given in equation (7), comparing the pre-treatment period 1996–2000 to the post-treatment period 2001–2003. Results are given in Table 6.

There is no evidence of an increase in the wages of highly educated trainees in response to the negative supply shock (column 2). This is in contrast to what standard economic theory predicts. To understand the absence of any wage effects, it is important to keep in mind that trainee wages are very rigid —often set by collective bargaining agreements— that the shock was only temporary, and that the supply of highly educated school graduates is fixed by the cohort size, giving very little scope for wage increases to increase their employment.[25]

[25] In fact, there are a multitude of reasons that potentially explain the lack of a wage adjustment. First, firms likely shy away from increasing wages in response to a temporary shock because downward rigid wages will impede a subsequent wage decline once the supply shock dissipates. Second, trainee wages in Germany are set at a very low level and are paid only throughout the three-year vocational training period. Hence, even a hypothetical doubling of training wages would result in negligible changes in absolute lifetime income. Instead, trainee supply responds to anticipated post-training wages (Neuber-Pohl et al., 2023) that remain unchanged in the present case. Third, the vast majority of training wages are set by collective bargaining agreements, and even firms that are not part of those agreements tend to base their wages on such agreements. Of course, firms could deviate upwards. In that case, worker's councils, which would have to approve training wages in large firms, would likely oppose unequal treatment of trainees. Finally, this finding is in line with the results by Muehlemann et al. (2022) in the case of the opposite, positive supply shock of trainees.

Table 6: DiD Results – Wage and worker substitution effects

	# highly educated trainees (1)	Log wage highly educ. trainees (2)	# low-educated trainees (3)	# highly educ. commuting trainees (4)	Log highly educ. VT employment (5)	Trainee retention rate (6)	Internal retraining (7)
Treat × Post	-1.01*	0.01	-0.94	2.08	0.07	-0.07	-0.04
	(0.47)	(0.03)	(1.42)	(2.66)	(0.07)	(0.04)	(0.06)
Mean dep. variable	4.14	2.99	11.65	5.19	2.25	0.59	0.47
N	3,133	2,025	3,133	1,364	2,930	2,989	1,552

Notes: DiD coefficients based on equation (7). Pre: 1997–2000. Post: 2001–2003. Standard errors clustered at the firm level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Among training firms only. Mean dep. variable: Average outcome in 1997–2000. 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 6: Based on the survey question "How many of the newly qualified apprentices are being offered a permanent position?". If missing, filled with the share of retained trainees from the administrative data. Column 7: Internal retraining is the number of retraining incidences per worker (fixed to pre-treatment levels). VT: completed vocational training.

Firms do not substitute for their missing highly educated trainees in any way. Prominent candidates as substitutes are low-educated trainees, highly educated trainees from other federal states, and highly educated workers who have already completed vocational training. However, firms do not compensate by hiring more low-educated trainees (column 3). In consequence, overall trainee employment also drops. The low substitutability between low- and highly educated trainees, in line with Muehlemann et al. (2022), is likely related to distinct skill sets, the specialization in different occupations, and the unchanged demand for low-educated trainees against a fixed supply of school graduates.

Also, there is no statistically significant increase in cross-state commuting of highly educated trainees from a different federal state following the shock (column 4).[26] No increased commuting supports the SUTVA assumption of no spill-overs across state borders.

Column 5 shows that the employment of highly educated workers who have completed their training program do not increase in response to the trainee shortage, indicating that already trained workers are not suitable substitutes for trainees.[27] Firms may try to compensate for missing trainees by retaining more trainees upon training graduation. Likewise, poaching of these workers might increase as well, such that the direction of the effect is ambiguous. I find no effect on the retention rate of recently graduated trainees (column 6). Firms may also increase retraining of incumbent workers to overcome skill shortages caused by the negative trainee supply shock. Again, I find no evidence for this (column 7).

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

[26] The coefficient of interest captures potential increases in commuting into treated states in addition to potentially reduced commuting into control states, and thus provides an upper bound of the true effect.

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

employment of workers with already completed vocational training.

7 Effects on firm technology investments

7.1 Effect on investments

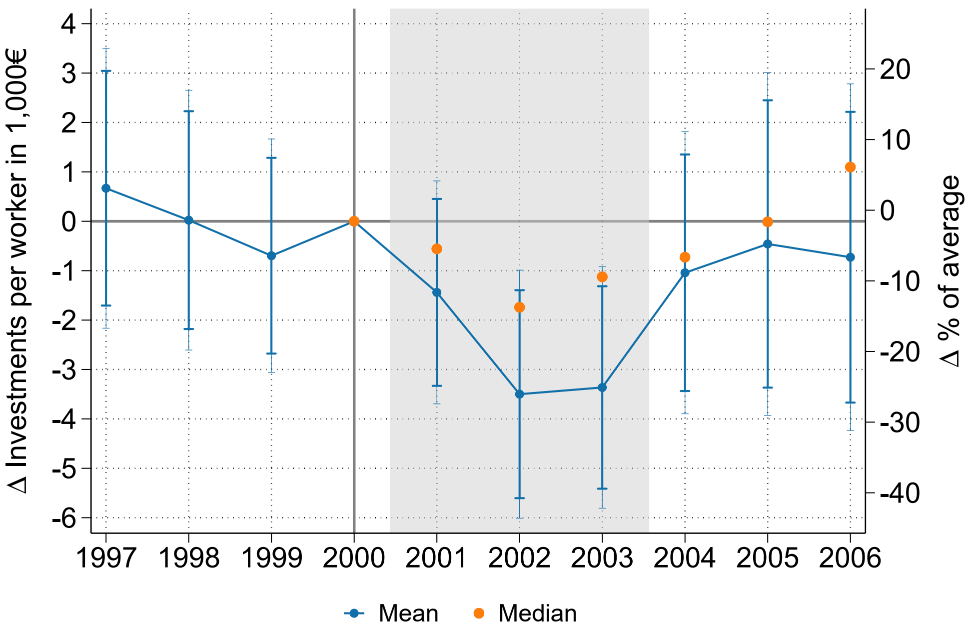
I now turn to the reform effects on firm investments, interpreting them as causal effects of facing a negative trainee supply shock. Figure 4 shows the difference in investments per worker between treated training firms and control training firms over time based on equation (6). Convincingly, there are no statistically significant pre-trends, supporting the assumption that investments in treated states would have evolved in parallel with investments in control states in the absence of the shock. The fact that firms do not adjust their investment behavior in anticipation of the trainee shortage, either upward or downward, may be due to two factors: First, firms are unaware of the impending shortage: According to the IAB establishment survey, in 2000, treated training firms did not expect significantly elevated shortages of young workers for 2001 and 2002 compared to control training firms. Second, opposing firm anticipatory behaviors may cancel each other: in the years before the shortage, some firms already forewent the adoption of new technologies, which require a critical mass of trainee cohorts to be trained on, because only few trainee cohorts will arrive in the years between reform announcement (1996, and 1998 respectively) and the trainee shortage. Other firms, however, invested more in anticipation, leading to net zero effects.

Firms time their investment to coincide with trainee supply: investments per worker significantly drop during the time of the trainee shortage in treated training firms compared to control training firms; the key finding of this paper. The decrease in investments is temporary: investments of treated firms go back to those of control firms in 2004–2006. This pattern clearly suggests that the decline is linked to the short-term drop in trainee supply, rather than a response to a new long-term reality in which trainees might be more highly skilled, older, or otherwise different.

Firms reduce investments by €2,740 per worker in 2001–2003, corresponding to a decline of 20% of pre-treatment average investments among treated firms, or 0.9% of the capital stock per worker. Given the right-skewed distribution of investments, I also compute the median across all firm-level effects. The median decrease, depicted by orange dots in Figure 4, in 2001 to 2003 is €1,140 per worker, equivalent to 8% of pre-treatment average investments per worker, or 0.4% of the capital stock per worker. The fact that the median declines as well clearly indicates that the effect is not driven by a few large outliers, but rather reflects a broad-based decline in investment across firms. To conclude, the estimated average decline in investments is large and goes beyond a potential "mechanical" effect of reducing capital in proportion to trainee employment. What can explain this large drop?

The first explanation relates to the temporary nature of the trainee supply shock: The

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



Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (6). Standard errors clustered at the firm level. Among training firms only. Left axis: Difference in investments in €1,000 divided by total employment in 1997 between treated and control firms. Right axis: Difference expressed in percent of the average pre-treatment investments per worker among treated firms. N=3,609. Mean: OLs estimates. Median: Median of the treatment effect distribution.

sizable estimate is plausible considering that the intertemporal elasticity of long-term investments is nearly infinite (see also House & Shapiro (2008)): firms face almost no cost from shifting investments by a couple of years, but gain substantially by postponing until trainees become available again. But if firms postpone investments, why do treated firms not compensate for their missing investments by investing more in the years after the shock compared to control firms? Several reasons could explain this. Investment frictions— organizational or financial—may cause treated firms to always lag behind control firms. Similarly, if investments trigger follow-up investments, a delay in the initial investment naturally causes delays in subsequent ones. Also, treated firms may decide to skip one technology vintage altogether, or a combination of these factors.

The second explanation for the large effect size relates to the realized distribution of trainees across firms in times of the shortage. Mechanically, investment effects are large if trainees are primarily missing in firms that would have made large investments in the absence of the shock, but much smaller or even zero, if only non-investing firms forego trainee employment. To assess this, I zoom in on the compliers of the shock, i.e. firms that decreased their employment of highly educated trainees in 2001–2003 relative to their pre-treatment levels, see Table 7. Indeed, complier firms are much larger than non-compliers, invest approximately 50% more, and have a 60% higher capital stock per worker, providing a plausible explanation for the large observed investment response. To abstract from the realized distribution of trainees, and to identify the effect of employing one trainee fewer, I use a different identification strategy below based on a Bartik-style instrument.

Table 7: Complier versus non-complier firms

	Complier	Non-complier	Δ
Δ # highly educated trainees	-2.24	+1.16	
# worker	283	205	77**
% highly educated workers with+in vocational training	11.6	9.2	2.4**
Investments per worker in €1,000	15.26	10.01	5.26***
Capital per worker in €1,000	390	238	152***
Share of firms	69%	31%	

Notes: Complying defined as decreasing average employment of highly educated trainees post compared to pre-treatment. Δ – Difference between compliers and non-compliers. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Δ highly educated trainees – Change in average employment of highly educated trainees post compared to pre-treatment

An alternative explanation for the large effect size is that treated firms reduce their overall employment in response to the shock, mechanically investing less if investments is proportional to employment. I can rule out this explanation: the estimated effect sizes are similar when using investment per worker based on time-varying firm size and when using investment per worker with firm size fixed at pre-treatment levels, see Table 8, columns 1 and 2.

Extensive versus intensive margin. Next, I study the extensive and intensive investment margin effect, see Table 8, columns 3 to 6. The extensive margin—measured as the binary outcome of investing versus not investing—is unaffected, which is expected since a very largy share of 92% of observations show positive investments. The intensive margin—measured as log investments—adjusts significantly.

The literature suggests that technology investment decisions are discrete rather than continuous since not all firms constantly adopt new technologies, and requires costly, indivisible ("lumpy") investments (e.g. Cooper et al., 1999; Bessen et al., 2020). In this setting, firms not planning to invest remain unaffected, while those intending to make lumpy investments may cancel due to the trainee shortage. To test this hypothesis, I look at a binary outcome of making large investments. Treated training firms are 10 percentage points (pp) less likely to make large investments (investments in the upper tercile of the investment per worker distribution) than control training firms when trainees are scarce, but are not differently likely to make very large investments (investments in the upper decile of the investment per worker distribution).

Table 8: DiD Results – Additional investment effects

	Inv. per worker in €1,000 constant workers	time-variant workers	Any inv. (1/0)	Log(Inv.)	Large inv. (1/0)	Very large inv. (1/0)
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	-2.74*	-2.63*	-0.01	-0.27*	-0.10**	-0.02
	(1.07)	(1.06)	(0.03)	(0.13)	(0.04)	(0.04)
Mean dep. variable	13.38	14.54	0.92	14.08	0.50	0.17
N	2,853	2,853	2,850	2,550	2,550	2,550

Notes: DiD coefficients based on equation (7). Pre: 1997–2000. Post: 2001–2003. Standard errors clustered at the firm level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Among training firms only. Mean dep. variable: Average outcome in 1997–2000. Large inv.: Investments in the upper tercile of the distribution of strictly positive investments per worker assigned as one, and zero otherwise. Very large inv.: Investments in the upper decile of the distribution of strictly positive investments per worker assigned as one, and zero otherwise.

Heterogeneity. I run separate regressions for firms in the skilled business service and public administration sector—where capital and labor tend to be complements—, and in the manufacturing sector—where capital and labor tend to be substitutes—, see Figure 5, Panel A. Interestingly, investments drop in both sectors, suggesting that a large variety of technologies adopted at the time are complements to new entrants.

One implicit assumption in the economic framework 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 50% of the trainees remaining at their training firms. However, there is variation across firms 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.

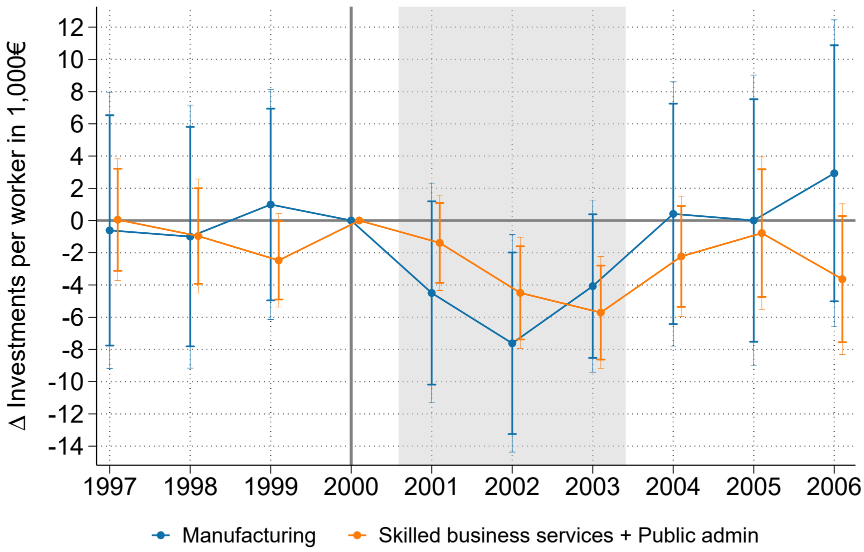
I run the event study regression including the triple interaction term Treat × Post × Trainee retention rate and all corresponding two-way interaction terms. The retention rate is defined as the pre-treatment proportion of trainees staying at the firm upon training completion.[28] Figure 5, Panel B, shows the predicted investment changes for firms with a high (low) trainee retention rate – evaluated at the 75th (25th) percentile of the distribution, corresponding to a retention rate of 68% (33%). Consistent with the hypothesized mechanism, firms with high retention rates reduce investments heavily in response to the reform, while treated firms with

[28]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.

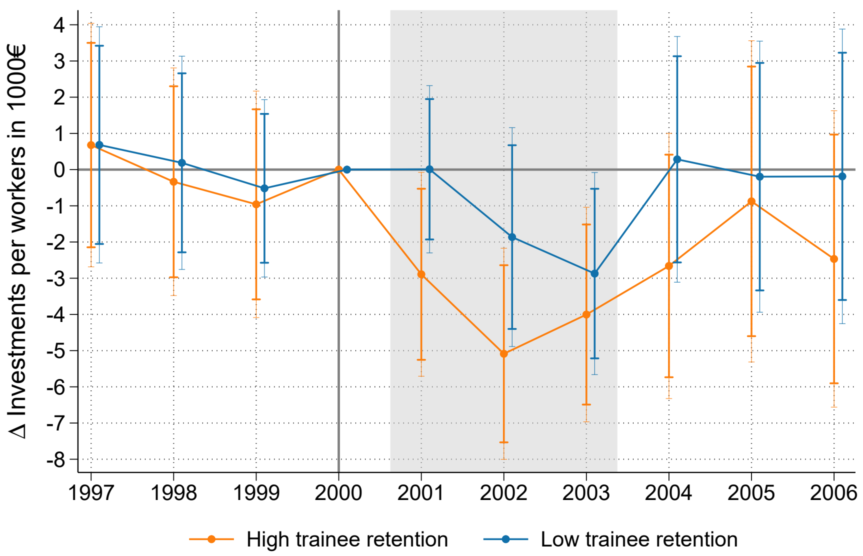
low retention rates reduce their investments much less.

Figure 5: Heterogeneity effects

A. Effects by industry



B. Effects by trainee retention rate



Notes: Panel A: based on a sample split with N=601 for manufacturing; N=2,088 for skilled business services and public administration. Panel B: based on the triple interaction effect Treat × Year × Trainee retention rate, evaluated at the 25th and 75th percentile of the trainee retention distribution.

Robustness. The negative effect of the reform on investments is robust to a wide range of specifications. To see this, I present the DiD estimates for investments per worker in Figure 6.

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

Firms may leave the sample over time. The negative estimate remains robust when restricting the sample to firms observed in every year between 1997 and 2003.

Convincingly, the effect is negative within both treated states, despite their differences in industry structure and geography. However, the effects are less precisely estimated due to the smaller sample size.

Even though I have shown that trainee migration or commuting across federal states seems to be limited, firms at federal state borders might be less affected by the reforms because they may attract trainees from control states. I use the share of commuters across federal states as a proxy for worker supply from other states. Excluding firms with a pre-treatment commuter share in the highest decile does not meaningfully affect the results.

When weighting the observations by the firms' pre-treatment employment size, the coefficient remains negative.

Acknowledging that states may be on different (linear) time trends, and controlling for them, does not meaningfully affect the results.

The results are also robust to different specifications of the matching procedure. When using the nearest neighbor instead of the three nearest neighbors, the estimated coefficient remains very similar, despite a significant reduction in sample size. The results are similarly robust to the inclusion of the 10% most distant matches.

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

Figure 6: Robustness

Notes: DiD coefficients and 90% and 95% confidence bands of the term Treat × Post. N indicates the number of observations in the respective estimation. *Control states:* Additionally including all West German training firms as control firms; dropping Berlin or Saxony from the set of control firms. *Balanced panel 1997-2003:* Sample restricted to firms observed in each year between 1997 and 2003. *Treated states separately:* Only using treated firms from one treated state and dropping firms from the other. *Excl. firms at border:* Dropping those 10% of firms with the highest 1999 cross-state commuter share of workers with vocational training. *Controlling for state trends:* Additionally controlling for linear state-specific time trends. *Matching procedure:* Using only the nearest neighbor instead of the three nearest neighbors as control firms; keeping all matches instead of discarding the furthest 10% of all matches.

under any permutation assignment, see Figure C3.

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 highest and lowest 2.5% (5%) of the draws under permuted treatment assignment are shown in Figure C2, Panel B. Again, the t-statistic of actual treatment assignment stands out as an outlier and is much smaller than the 5% and 2.5% most negative t-statistics under permuted treatment assignment.

Falsification test among non-training firms. To validate that the investment decline is indeed linked to the trainee shortage and not due to some idiosyncratic factor happening in the two treated states around this time, I turn to the sample of non-training firms, i.e. firms operating in the same industries but with no highly educated trainees throughout 1995–1997. Non-training firms should be much less affected by the reform, at most because they would want to start training but cannot during the time of the trainee shortage, or via spillover effects. Table 9 shows the event study estimates for both training and non-training firms.

Figure 7: Permutation test – T-statistics

Notes: T-statistics of the event study regression coefficients based on equation (6) using the actual treatment assignment (orange line) and all possible permutation assignments within East Germany (gray lines).

Looking at column 1, to a small extent, non-training firms in control states started employing highly educated trainees in 2001–2003, though the resulting negative treatment effect is only a tenth of the effect among training firms. While the investment drop among non-training firms is not exactly zero, likely reflecting spillover effects or the small decrease in highly educated trainee employment, we see much larger declines in investment among training than among non-training firms. This provides further evidence that the investment drop is indeed related to the negative trainee supply shock.

Table 9: Training versus non-training firms

	A. Training firms			B. Non-training firms		
	# highly educated trainees	Inv. per worker in €1,000	Log(inv.)	# highly educated trainees	Inv. per worker in €1,000	Log(inv.)
	(1)	(2)	(3)	(4)	(5)	(6)
Treat × Post	-1.01*	-2.74*	-0.27*	-0.09***	-1.04*	-0.13
	(0.47)	(1.07)	(0.13)	(0.03)	(0.53)	(0.09)
N	3,133	2,853	2,550	9,683	8,832	6,706

Notes: DiD coefficients based on equation (7). Pre: 1997–2000. Post: 2001–2003. Standard errors clustered at the firm level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 10: IV results – Second stage

	Inv. per worker in €1,000 (1)	Log(investments) (2)
$N^{Trainee}$	0.73**	0.06**
	(0.30)	(0.03)
F-Stat	12.27	14.70
N	6,686	5,007

Notes: F-Stat gives the robust Kleibergen-Paap Wald rk F statistic. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

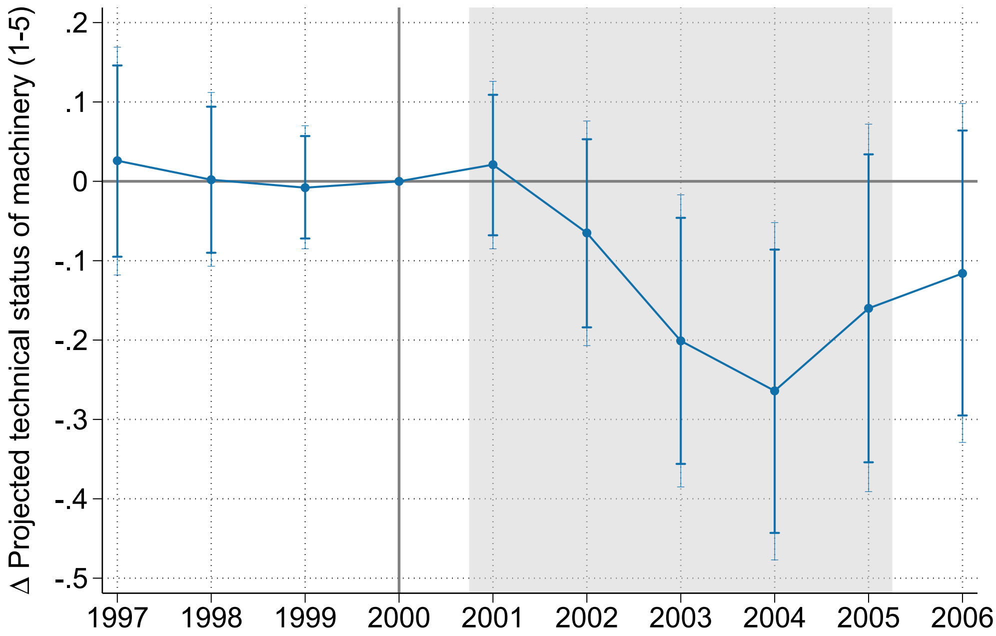
Firm-level treatment intensity – Instrumental variable regression. Another way to strengthen the argument that the investment decline is indeed caused by the negative trainee supply shock is by analyzing whether firms that suffer from larger reform-induced trainee employment reduce investments more. To do so, in a complementary analysis, I instrument firms' trainee employment with a Bartik-style instrument based on firms' initial employment of highly educated trainees (i.e. exposure to the reform; *share*) and the reform (i.e. *shift*). I extensively discuss the identification strategy and report results in Appendix D. This analysis not helps tying the trainee shortage and the investment decline, it also provides an estimate of the investment decline associated with each absent highly educated trainee. Also, in the above DiD event study design, the identified coefficient is subject to the realized distribution of trainees across training firms, where I have found that compliers tend to be large, heavily investing firms. The IV strategy explicitly takes this selection into account.

The IV analysis reveals that more exposed training firms indeed experience larger employment decreases of highly educated trainees and reduce investments more. In particular, each missing highly educated trainee reduces firm investments by approximately €730 per worker, or 0.05log points, see Table 10. This figure is substantially lower than the one implied by the ratio between missing trainees and investments identified in the event study regression above. This may be due to spill-over effects across firms within treated states, due to reduced trainee quality during the time of the shortage, or the fact that firms complying with the DiD identification strategy tend to be large, heavily investing firms.

7.2 Effect on firm technology adoption

Having established that the reform-induced trainee shortage decreases firm capital investments, the following section investigates whether this is linked to foregone technology adoption. In particular, I look at the self-assessed technical status of a firm's machinery on a continuous scale from 1 ('completely out-of-date.') to 5 ('state-of-the-art'), projected back to the years when investments affecting it took place, see again Section 4. As shown in Figure 8, treated training firms report an outdated projected technical status of their machinery compared to control firms starting in 2002. The depreciation is meaningful in magnitude: a depreciation by approximately -0.26 in 2004 corresponds to 26% of firms reporting a deterioration by one category, or a decrease by a third of a standard deviation. Treated firms' technical status approximately converges back to control firms' technical status by 2006, indicating that treated firms' foregone investments do not put these firms on a different long-term trend. In contrast, treated firms manage to catch up with control firms once trainees are available again, or simply skip one technology vintage.

Figure 8: Effect on technology adoption



Notes: Event study coefficients of the interaction terms Treat × Year plus 90% and 95% confidence bands. Based on equation (6). Standard errors clustered at the firm level. Among training firms only. Projected technical status: Technical status of a firm's machinery on a scale from 1 ('completely out-of-date.') to 5 ('state-of-the-art'), projected to the year in which the investments influencing that status were made. N=3,025

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

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

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 suggests that young labor market entrants are complements to firm technology adoption which I explain by their low opportunity costs and/or high expected returns from skill acquisition. Consequently, in times when young labor market entrants are scarce, firms face higher capital adjustment costs of worker training, which potentially renders the adoption of technologies that require new skills unprofitable. This complementary relationship is likely not limited to the context of German vocational trainees, as the mechanism is likely to hold across a range of settings, and since young workers are consistently found to work more often with new technologies across education levels and countries.

While it has been known that labor supply affects technology adoption, this paper introduces the crucial role of young labor market entrants and the mechanism via capital adjustment costs of training new skills. These insights are novel and informative from a number of perspectives: First, they highlight the relevance of new skills demanded by new technologies, making the comparative advantage in skill acquisition an important factor determining complementarity with new technologies—beyond the factors previously emphasized, namely skill and task levels. Second, the results highlight that the availability of new entrants is a key determinant of firm technology adoption. While a reduction in the supply of young labor market entrants likely not always leads to a decrease in technology adoption, it always raises its cost. This insight is relevant given current shortages of young workers in most developed countries. Third, the findings imply that retraining incumbents is costly compared to training young labor market entrants. This leads to strong vintage effects, where different worker cohorts posses distinct, vintage-specific skills, which slows down technological transitions (Adão et al., 2024).

While studying a short-lived labor supply shock, by establishing complementarity between new entrants and new technologies, this paper also informs on the macro debate on endogenous technological change and demographic change: countries with lower population growth or shortages of middle-aged workers are found to adopt more (labor-saving) robots (Abeliansky & Prettner, 2017; Acemoglu & Restrepo, 2022). The "new-skills" channel, which this paper is able to separate out, works in the other direction as the labor-saving channel, calling for a more nuanced view of the effect of demographic change on firm technology adoption. In line with the "new-skills" argument, Angelini (2023) finds that above a certain tipping point, population aging reduces investments in information and communication technologies.

From a policy perspective, the findings stress the importance of attracting young labor market entrants or subsidizing the retraining of incumbent workers to foster technology adoption. The results also have implications for the optimal design of education systems: the finding that firms shy away from retraining incumbent workers who were trained a few years ago indicates that skills acquired through vocational training may be overly specific (compare Hanushek et al., 2017).

References

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

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

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

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

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

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

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

Adão, R., Beraja, M., & Pandalai-Nayar, N. (2024). Fast and slow technological transitions. Journal of Political Economy Macroeconomics, 2(2), 000–000.

Aghion, P., Akcigit, U., Hyytinen, A., & Toivanen, O. (2024). A year older, a year wiser (and farther from frontier): Invention rents and human capital depreciation. Review of Economics and Statistics, 106(4), 974–982.

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

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

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

Arntz, M., Lipowski, C., Neidhöfer, G., & Zierahn-Weilage, U. (2025). Computers as stepping stones? technological change and equality of labor market opportunities. Journal of Labor Economics, 43(2), 000–000.

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

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

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

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

Barsi, M. (2001). Betriebsbesuch: Arbeitskräfte reichen nicht – landtechnik gmbh sucht noch fachleute. Mitteldeutsche Zeitung. (Ausgabe: Weissenfels)

Bartel, A. P., & Sicherman, N. (1998). Technological change and the skill acquisition of young workers. Journal of Labor Economics, 16(4), 718–755.

Barth, E., Davis, J. C., Freeman, R. B., & McElheran, K. (2023). Twisting the demand curve: Digitalization and the older workforce. Journal of Econometrics, 233(2), 443–467.

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.

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

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

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

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.

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.

Danzer, A. M., Feuerbaum, C., & Gaessler, F. (2024). Labor supply and automation innovation: Evidence from an allocation policy. Journal of Public Economics, 235, 105136.

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

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

Dorner, M., Görlitz, K., & Jahn, E. J. (2024). The impact of a missing school graduation cohort on the training market. Economics of Education Review, 103, 102580.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Fitzenberger, B., Holleitner, J., & Kagerl, C. (2024). Herausforderungen für die arbeitsmärkte der zukunft am beispiel deutschland. Wirtschaftsdienst, 104(8), 519–523.

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

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

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

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

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

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

House, C. L., & Shapiro, M. D. (2008). Temporary investment tax incentives: Theory with evidence from bonus depreciation. American Economic Review, 98(3), 737–768.

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

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

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

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.

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). Expertise at work: New technologies, new skills, and worker impacts. ZEW-Centre for European Economic Research Discussion Paper, 24-044.

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

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

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.

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

Roth, J., Sant'Anna, P. H., Bilinski, A., & Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. Journal of Econometrics, 235(2), 2218–2244.

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

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

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.).

Siegloch, S., Wehrhöfer, N., & Etzel, T. (2025). Spillover, efficiency, and equity effects of regional firm subsidies. American Economic Journal: Economic Policy, 17(1), 144–180.

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.

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

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

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

A Data

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

Variable	Survey Question	Manipulation	Frequency
Inv. per worker	What was the approximate sum of all investments in t?	Divided by number of workers in 1997 from the administrative records. Trimming uppest percentile of the investment distribution and the investment per worker distribution	Yearly
Inv. type (0/1)	Did your establishment invest in one or more of the following areas in the last business year of t? EDP, information and communication technology? Production facilities, plant and equipment, furniture and fixture? Means of transport, transportation systems? Real estate and buildings?		Yearly
Technical status of machinery	How do you assess the overall technical status of the plant and machinery, furniture and fixtures of this establishment compared to other establishments in the same industry? "1" - state-of-the-art equipment. "5" - completely out-of-date.	Inverted order	Yearly except for 2004

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

B Potential reform confounders

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

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

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

Second, Germany's main regional policy, GRW, aimed at revitalizing underdeveloped regions, particularly East Germany, through investment subsidies for (mainly) manufacturing plants. The maximum subsidy rate varied based on counties' economic performance indicators, and was frequently reformed between 1997 and 2014. Siegloch et al. (2025) exploit these

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

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

Notes: Panel A: Treated: Mecklenburg-Western Pomerania and Saxony-Anhalt. Post: 2001 onward. *Panel B:* Social democrats among governing parties (1/0). Controlling for state and year fixed effects. Observations at the state-year level for East German states for 1992 until 2005, except for education expenditure (column 2 and 3) that is only observed from 1995 onward. Education expenditure: Total public expenditure on education. Share education expenditure: Public expenditure on education as a percentage of the total budget. Unemployment rate: Unemployment rate in % of the dependent civilian labor force. Debt: Debt of the overall public budget. Sources: (1) – Federal Statistical Office (2022) (2) & (3) – Federal Statistical Office (2023b) (4) – Federal Statistical Office (2023a) (5) – Federal Statistical Office (2023e) (6) – Federal Statistical Office (2023d) (7) – Federal Statistical Office (2023c) * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

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

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

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

C Additional results

Table C1: Technical status of machinery and lagged firm investments

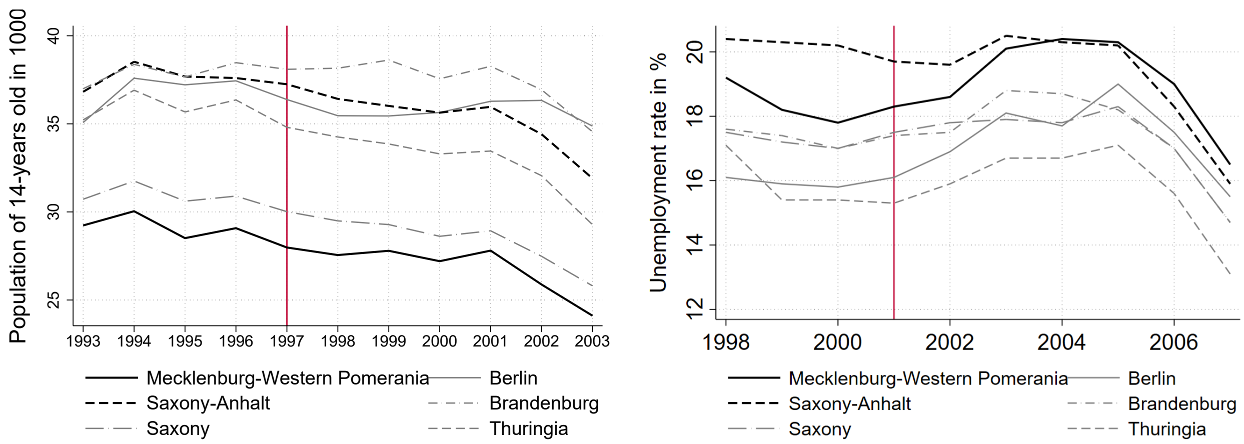
Inv. per worker	0.0010
	(0.0005)
Inv. per worker $t-1$	0.0053***
	(0.0005)
Inv. per worker $t-2$	0.0033***
	(0.0005)
Inv. per worker $t-3$	0.0017***
	(0.0005)
Inv. per worker $t-4$	0.0013**
	(0.0005)
Inv. per worker $t-5$	0.0020***
	(0.0005)
N	28,279

Notes: Outcome: Technical status of machinery in t. Controlling for year and firm fixed effects. Standard errors clustered at the firm level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Figure C1: Demographic and economic trends across federal states

A. Number of 14-years-old by state

B. Unemployment rate by state



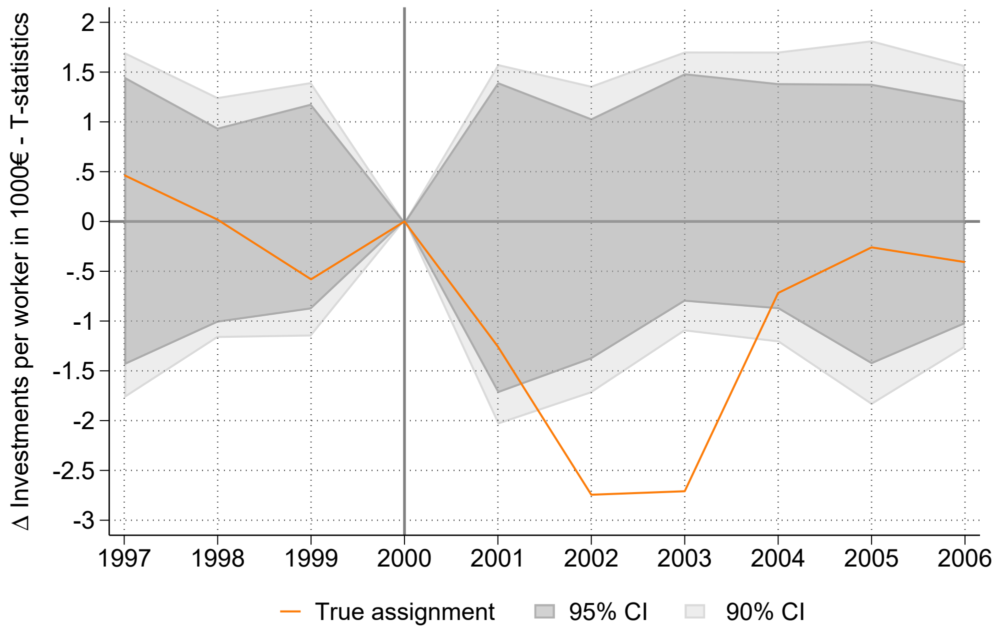
Notes: Panel A: Source: Federal Statistical Office, Genesis-Online (2022b). The number for Saxony is divided by two for better visibility. *Panel B:* Source: Federal Statistical Office, Genesis-Online (2022b).

Table C2: Correlation of the error term within states and within firms

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

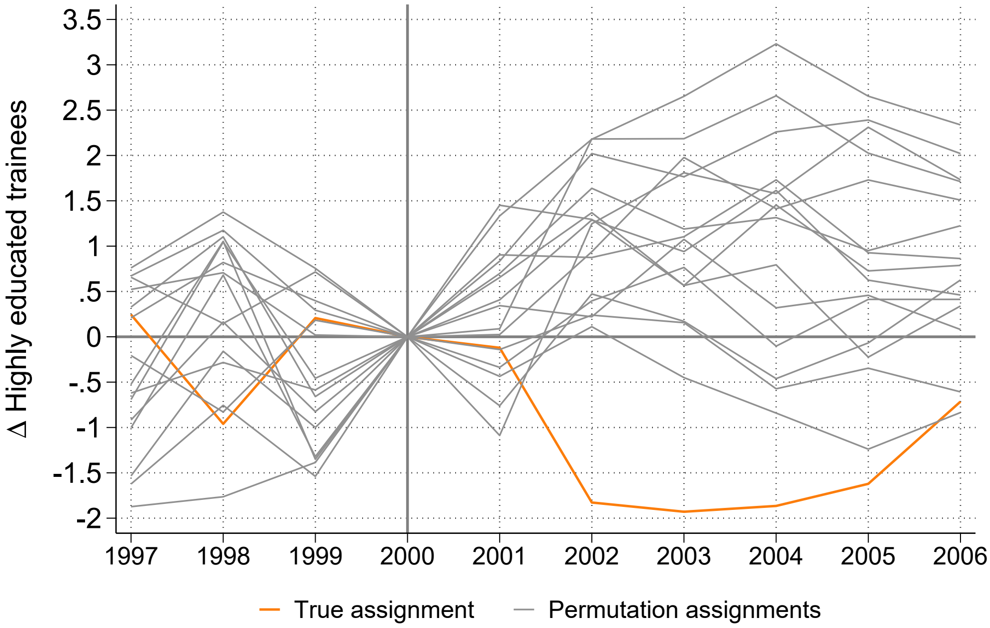
Notes: Standard deviations (SD) of the error term resulting from a regression following equation (6) with the outcome variable shown in the first column.

Figure C2: Permutation test in West Germany – T-statistics; Effect on investments



Notes: T-statistics of the DiD event study coefficients based on equation (6) using the actual treatment assignment (orange line) and all possible permutation assignments across West Germany (gray areas).

Figure C3: Permutation test in East Germany – T-statistics; Effect on trainee employment



Notes: T-statistics of the DiD event study coefficients based on equation (6) using the actual treatment assignment (orange line) and all possible permutation assignments across East Germany (gray lines).

D Instrumental variable regression

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

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

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

$$\begin{aligned} N_{jt}^{\text{Trainee}} &= \sum_t \gamma_t (N_{j,1996/97}^{\text{Trainee}} \times \text{Treat}_j \times \text{Year}_t) \\ &\quad + \sum_t \zeta_t (N_{j,1996/97}^{\text{Trainee}} \times \text{Year}_t) + \psi_t + \pi_j + \epsilon_{jt} \end{aligned} \tag{D2}$$

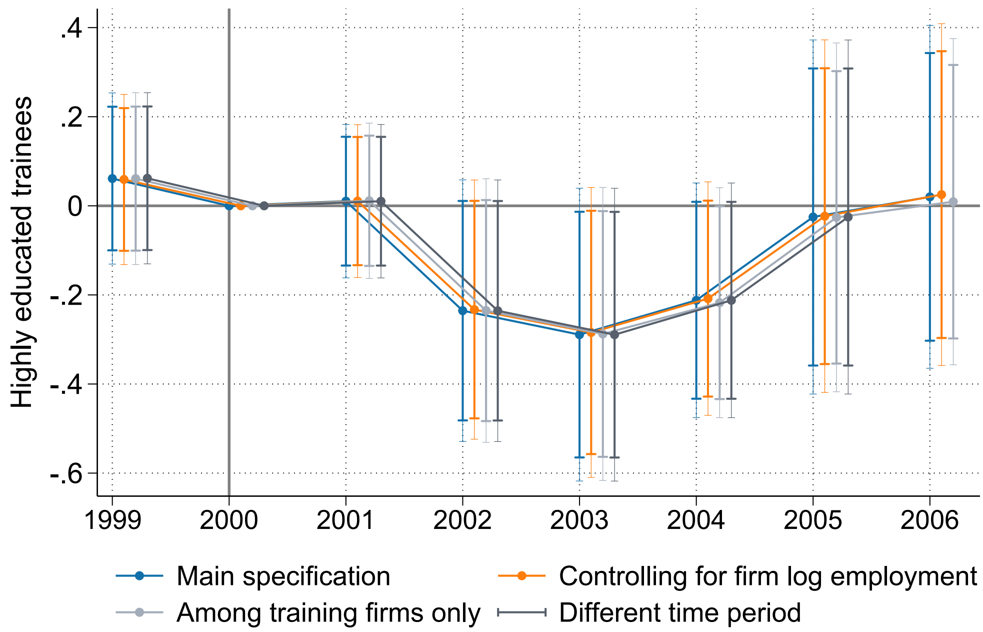
with j firms, and t calendar years. Treat takes the value one if a firm is located in a state undergoing the education reform and zero otherwise. I predict contemporaneous trainee employment by firms' initial mean employment of highly educated trainees in 1996/1997, i.e. firm exposure, $N_{1996/97}^{\text{Trainee}}$, corresponding to the *shares* in a shift-share instrument, times Treat $\times$ Year, corresponding to the reform-induced *shifts* in the supply of trainees across states and years. I control for time trends in firm trainee employment which are allowed to vary by firm exposure, $N_{j,1996/97}^{\text{Trainee}} \times \text{Year}_t$. Hence, the instrument exploits variation between two equally exposed firms located in a treated state and a control state across time. The exogeneity of the instrument stems from the random assignment of the trainee supply shock, i.e. the education reform, to states and years. Since employment of highly educated trainees in 1996/97 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.[29] I estimate the effect within the sample of matched firms,[30] ensuring that treated and control firms are comparable in terms of sector, size, and employment of highly educated trainees.

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

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

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

Figure D1: IV results – First stage



Notes: Coefficients plus 90% and 95% confidence intervals of the term ($N^{\text{Trainee}}_{j,1996/97} \times \text{Treat}_j \times \text{Year}_t$) in equation (D2). Outcome: Number of highly educated trainees. Standard errors clustered at the firm level.

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

Table D1, Panel A, also shows the results of the second stage for different specifications of the investment outcome. While only borderline significant, the positive coefficients imply that a reduction in trainee employment decreases firm investments. In particular, one fewer trainee reduces investments by €930 per worker (column 1), equivalent to 7% of average investments. The investment decline is driven by the intensive investment margin (column 3), in line with the results based on the DiD event study. Note that the estimate is identified for reductions in trainee employment of around 1.1 (as we know from the event study) but is unlikely to be linearly scalable for substantially larger drops.

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

When restricting the sample to training firms only (Panel C), the estimates look comparably, but are mostly not statistically significant. I conclude that the distinction between employing or not employing trainees is more relevant for firm investments than the number of trainees conditional on having at least one trainee.

Table D1: IV results – Second stage

	(1)	(2)	(3)	(4)
		A. Investments per worker in €1,000		
$N^{Trainee}$	0.73**	0.71**	0.39*	0.77***
	(0.30)	(0.30)	(0.22)	(0.28)
F-Stat	12.27	12.53	11.91	13.89
N	6,686	6,686	1,708	6,170
		B. Log(investments)		
$N^{Trainee}$	0.06**	0.05*	0.04	0.05*
	(0.03)	(0.03)	(0.03)	(0.03)
F-Stat	14.70	15.19	12.35	16.17
N	5,007	5,007	1,494	4,608
Time period	1999-2006	1999-2006	1999-2006	1999-2005
Controls		X		
Firm sample	All	All	Training only	All

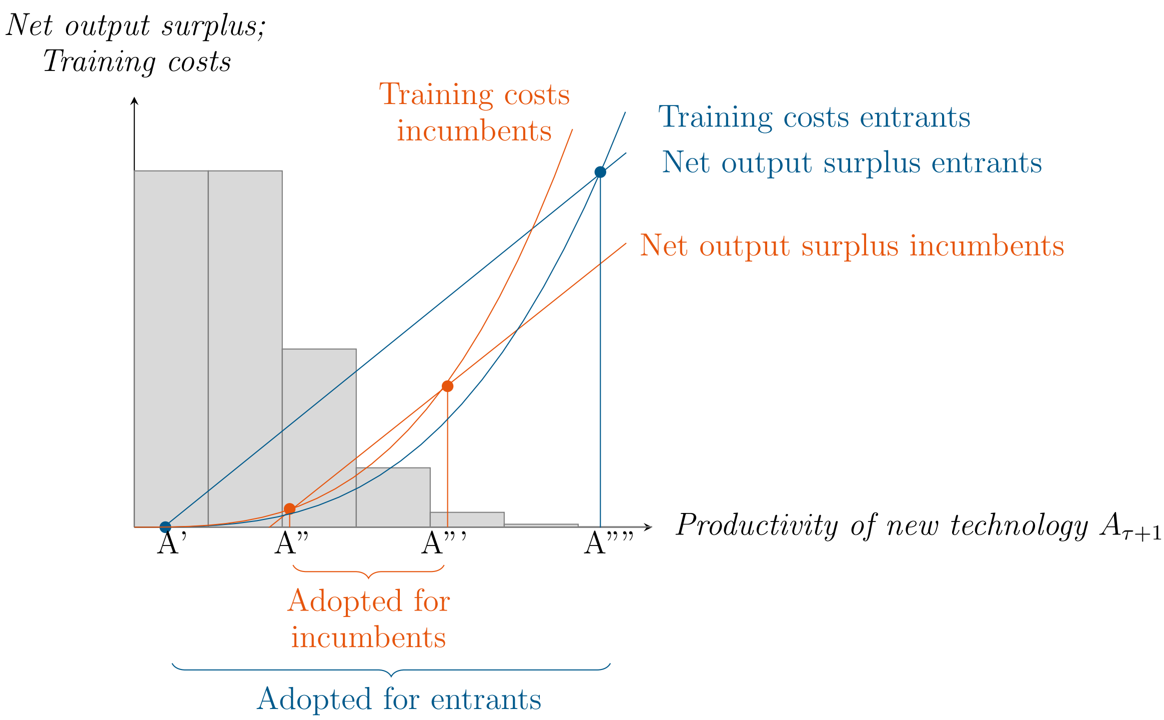
Notes: F-Stat gives the robust Kleibergen-Paap Wald rk F statistic. Based on the matched sample. Inv. per worker: investments in €1,000 divided by total employment in 1997. Standard errors clustered at the firm level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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

E Economic framework

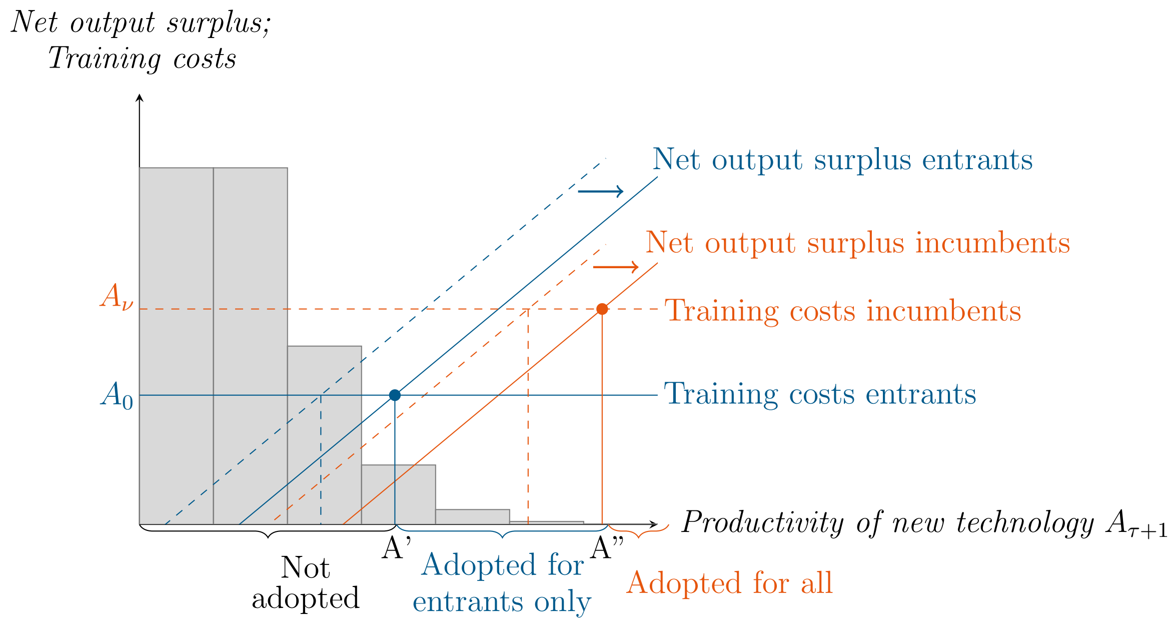
Extension – Increasing and convex capital adjustment costs. Until now, I have assumed training costs to be constant over all productivity levels. In standard capital adjustment costs models, adjustment costs are assumed to be increasing and convex in investment size. Let us now assume that training costs are increasing and convex in technology productivity, $C'(A_\nu) > 0, C''(A_\nu) > 0$. This implies that small investments can be easily incorporated in the structure of the firm without much training, while large investments create more pronounced disruptions requiring longer training. The trade-off between additional profits and additional costs of production is shown in Figure E1 and is similar to the model in Acemoglu & Pischke (1999). New technologies below a productivity threshold A' are not adopted because training costs are too high; new technologies above A' but below A'' are adopted by training labor market entrants only; and new technologies above a threshold A'' are adopted by retraining incumbent workers as well. In addition, there are new, highly productive technologies above a certain productivity threshold A''' that require prohibitively long training to justify retraining incumbents, and even more productive technologies above a certain productivity threshold A'''' for which training costs exceed productivity gains for all workers. Consequently, a lack of entrants not only hinders the adoption of rather unproductive new technologies in the range between A' and A'', but also of very productive technologies with productivities in the range between A''' and A''''.

Figure E1: Firms' costs and benefits of technology adoption with convex adjustment costs



Notes: Profitability of assigning new technology-using task to entrants versus incumbent workers when capital adjustment costs of training are increasing and convex in technology productivity. The histogram shows the productivity distribution of the new technology vintage $A_{\tau+1}$.

Figure E2: Firms' costs and benefits of technology adoption with worker retention rates below one



Notes: Profitability of assigning new technology-using task to entrants versus incumbent workers. The histogram shows the productivity distribution of the new technology vintage $A_{\tau+1}$. Dashed lines show net output surplus assuming full worker retention. Solid lines show net output surplus assuming worker retention rates below one.

F Supporting evidence based on additional datasets

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

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

IAB/BIBB/BAuA Qualification and Career Surveys. This analysis is based on the 1999–2012 waves of the IAB/BIBB/BAuA Qualification and Career Survey (QCS). The QCS are repeated cross sectional surveys conducted by BIBB, IAB, and BAuA.[31] The survey covers around 30,000 employees. With the only difference not excluding trainees and East Germany, I closely follow Arntz et al. (2025) and I use the harmonization and data restriction procedure by Rohrbach-Schmidt & Tiemann (2013).

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

European Working Conditions Survey. I use the 2000 wave of the European Working Conditions Survey (EWCS), in which workers report how often their main job involves working with computers, laptops, or smartphones on a 7-point scale ranging from "all of the time" to "never". From this, I construct a binary outcome variable and regress it on a dummy variable indicating whether the worker is under 30 years old, and a range of detailed covariates, including occupation and sector, as above. I restrict the sample to the EU-15 countries (those countries that were members of the European Union before the 2004 enlargement) and exclude self-employed workers. The results are shown in Table E1: Being young significantly increases computer use at work in 2000— by between 2.4 percentage points (6%) and 3.5 percentage points (14%), depending on the specification of the outcome variable.

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

Table E1: Computer use by age group across EU-15

	(1)	(2)	(3)
$<$ 30 years	0.024*	0.035**	0.025*
	(0.012)	(0.011)	(0.011)
Mean dep. variable	0.43	0.25	0.20
N	17,056	17,056	17,056

Notes: Based on the European Working Conditions Survey in 2000, including Belgium, Denmark, Germany, Greece, Spain, France, Ireland, Italy, Luxembourg, Netherlands, Austria, Portugal, Finland, Sweden, and United Kingdom. Based on the survey question "Does your main job involve working with computers, laptops, or smartphones? – (1) All of the time, (2) Almost all of the time, (3) Around 3/4 of the time, (4) Around half of the time, (5) Around 1/4 of the time, (6) Almost never, (7) Never. Column 1: Computer use equal to one for (1)–(5); and zero otherwise. Column 2: Computer use equal to one for (1)–(3); and zero otherwise. Column 3: Computer use equal to one for (1)–(2); and zero otherwise. Controlling for country fixed effects, broad industry categories (4), gender, public sector employment, country-specific occupation fixed effects, and tenure at firm in three-year bins. Heteroscedasticity-robust standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.