

DISCUSSION PAPER SERIES

IZA DP No. 16652

**How Negative Labor Supply Shocks
Affect Training in Firms: Lessons from
Opening the Swiss-German Border**

Caroline Neuber-Pohl
Damiano Pregaldini
Uschi Backes-Gellner
Sandra Dummert
Harald Pfeifer

DECEMBER 2023

DISCUSSION PAPER SERIES

IZA DP No. 16652

How Negative Labor Supply Shocks Affect Training in Firms: Lessons from Opening the Swiss-German Border

Caroline Neuber-Pohl

BIBB

Damiano Pregaldini

University of Zurich

Uschi Backes-Gellner

*Stanford University, University of Zurich
and IZA*

Sandra Dummert

IAB

Harald Pfeifer

BIBB, ROA, Maastricht University and IZA

DECEMBER 2023

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

How Negative Labor Supply Shocks Affect Training in Firms: Lessons from Opening the Swiss-German Border*

By exploiting a labor market reform causing an outflow of German workers to Switzerland, we examine the effect of negative labor supply shocks on training in firms using the market for apprenticeships as an example. Analysis of administrative data reveals that the reform led to more apprentices in German firms despite a decrease in apprentice wages. This can be explained by a standard two-factor production model where firms substitute outflowing skilled workers with more apprentices; setting lower wages is possible because of a rising supply of apprentices owing to substantially improved employment prospects after border openings.

JEL Classification: J21, J22, J61, R23

Keywords: negative labor supply shock, training effects after worker outflow, wage effects, training incentives, apprenticeship training supply and demand

Corresponding author:

Uschi Backes-Gellner
Stanford University
450 Jane Stanford Way
Stanford, CA 94305
USA
E-mail: ubg@stanford.edu

* We thank the participants at the First International Conference on the Economics of Vocational Education and Training 2022, COPE 2023, IWAAE 2023, IAAE 2023, and SASE 2023 for their invaluable feedback. We are particularly grateful for the very constructive comments from Eric Bettinger and Stefan Wolter. Additionally, Uschi Backes-Gellner extends her gratitude to Stanford University's Graduate School of Education for its hospitality and supportive environment during her visiting scholarship. We also thank the Swiss State Secretariat for Education, Research and Innovation for funding through its "Leading House VPET-ECON: A Research Center on the Economics of Education, Firm Behavior and Training Policies."

I. Introduction

This study answers the question of how negative labor supply shocks affect local adjustments of training in firms. We exploit the free movement of labor reform implemented in 2002, which allowed for sudden access of German workers to the Swiss labor market. Taking advantage of the substantially higher wages in Switzerland, German workers commuted in large numbers from the German border regions to Switzerland after the reform. This resulted in a sizable exogenous negative labor supply shock on the German side of the border.

While previous studies have examined the effect of such immigration shocks on the employment and wages of workers in the receiving (e.g., An et al., 2022; Beerli et al., 2021; Borjas, 2003; Card, 1990, 2001; Dustmann et al., 2005; Dustmann and Glitz, 2015; Dustmann et al., 2016, 2017; Glitz, 2012; Ottaviano and Peri, 2012) or sending (Bütikofer et al., 2019; Dicarolo, 2022; Hafner, 2022) country, we examine for the first time training adjustments in firms in the sending country using the German apprenticeship system as an example. In Germany, apprenticeships have historically been an instrument for firms to meet their skill demands. However, as in many other countries and training systems, the German apprenticeship system has recently come under pressure because of an insufficient supply of young adolescents applying for apprenticeship positions. In light of this development, it is important to analyze how training markets function and how training decisions of individuals and firms are affected by negative labor supply shocks.

We derive the hypotheses on the direction of expected changes in apprenticeship numbers and wages after the opening of borders from a textbook short-run production model with only two types of labor inputs: first, skilled workers who have already completed an apprenticeship¹ and second, apprentices. The model assumes a competitive labor market and that the two types of labor are highly substitutable because apprentices can already (partially) perform the tasks of skilled workers during

¹In the following, we refer to apprenticeship graduates as skilled workers not including university graduates.

their training period. We show that the model simultaneously predicts an increase in the employment of apprentices and lower wages for apprentices.

For our empirical analysis, we apply a difference-in-differences (DiD) framework leveraging two sources of exogenous variation in the intensity of treatment. First, we exploit the variation over time owing to a step-by-step introduction of the border opening, which gradually lifted the restrictions for cross-border commuters in two steps in 2002 and 2004. Second, we exploit variation in treatment intensity by geolocation due to varying access to border crossings. We use rich administrative data sets from both sides of the border, Germany and Switzerland. These are ideal for our purpose because they allow us to zoom in on the affected border regions and identify high- and low- treated firms using their location, as we could use highly disaggregated geolocated travel distances to the Swiss labor market. This identification strategy is effective because the data show that the opening of borders for German workers almost only affected commuting behavior in municipalities with very short travel distances to Switzerland (less than 30 minutes). In addition to changes over time, we could account for the differences in treatment intensity within the regions affected by the reform. To isolate the effect of the negative labour supply shock on apprenticeship training, we compare the changes in training firms in treated regions close to the border and in treated regions located farther away from the border with firms in matched control regions located outside the regions affected by the reform. We exploit both sources of variation in treatment intensity (over time and by location) to study the causal effect of the reform on training.

Our empirical results show that, as predicted by our theoretical model, the number of apprenticeships increased in the sending country because of the negative labor supply shock despite a decrease in apprentice wages that occurred simultaneously. Training in firms increased at both extensive (firms providing training) and intensive (number of apprentices) margins. Overall, we not only show how training in firms changes as a consequence of a negative supply shock owing to opening borders to a richer country but also provide an ideal empirical example and use case for a textbook

application of a short-term two-factor production model.

This study contributes to the literature in three ways. First, it complements the literature on the effects of labor migration between industrialized countries (Bütikofer et al., 2019; Dicarolo, 2022; Hafner, 2022). While previous studies have focused on the effects of worker inflows on native employment and wages in the sending country and the effects of worker outflows on employment and wages, none have studied the effects of worker outflows on training in firms. We close this gap by studying the effects of training on firms in the sending countries and ultimately expound on the effects of an outflow of workers on longer-term regional labor market prospects and human capital formation in the sending country after the opening of borders.

Second, we add to the classic literature on labor demand and supply models by presenting a textbook application of input adjustments in tight labor markets. Building on the simple two-factor production model proposed by Borjas (2003), we show that, in the presence of skilled worker shortages, apprentices serve as substitutes for skilled workers if the longer-term wage prospects of skilled workers are attractive enough to secure an increased supply of apprentices. The upward adjustments of apprentice training in firms following the negative supply shock of skilled workers align with the classic theory of labor market adjustments.

Third, by concentrating on the effects of apprenticeship training, our study adds to the broader literature on the effects of labor market shocks on the market for apprenticeships (Aepli and Kuhn, 2021; Dorner and Görlitz, 2020; Lüthi and Wolter, 2020; Muehlemann et al., 2022; Oswald-Egg and Siegenthaler, 2021; Wittek and Muehlemann, 2021). By combining the clean identification of an emigration-driven negative labor supply shock on affected workers and firms, we can draw important conclusions regarding the functioning of apprenticeship training markets. On the supply side of apprentices, the attractiveness of apprenticeship training evidently depends more on the long-term wage prospects of apprenticeship graduates, that is, on attractive skilled worker wages than on short-term apprentice wages. In our empirical example, firms can hire more apprentices despite (moderately) lower training wages because the sup-

ply of apprentices increases in response to the higher wages that skilled workers gain after the opening of the borders. On the demand side, firms also react to the negative supply shock and the subsequent change in relative labor costs by substituting apprentices for skilled workers and reducing apprentice wages to compensate for the higher risk of losing their skilled workers to the other side of the border. This suggests that firms are sensitive to both the costs of skilled workers and the (short-run) costs of apprenticeships.

While our results concern the market for apprenticeship training in Germany, they expand to other countries with apprenticeship training systems and to other types of occupations in which training takes place while working in companies or public institutions such as training of PhDs, auditors, lawyers, or teachers. Our results generally imply that firms are willing and able to use the training of young labor market entrants as a means to meet their skill demands if the cost-benefit ratio of training is advantageous. However, the key to mobilizing the young generation to invest in training is to increase the long-term benefits of training, that is, the wages that graduates of apprenticeships receive (and the societal reputation accompanying it).

II. Negative labor supply shocks in a two-factor model

To develop hypotheses on the possible short-run effects of a negative labor supply shock, we apply a textbook two-factor production model as presented by Borjas (2003) and recently applied to a positive labor supply shock on the market of apprenticeships by Muehlemann et al. (2022). We consider two types of labor: 1. skilled workers who have already completed an apprenticeship and 2. apprentices. Skilled workers x_{SW} require a vocational degree; in our empirical context, that is, Germany, it exists in the form of an apprenticeship. Apprentices x_{AP} have not yet undergone training but are productive at the workplace during the training period and are fully productive thereafter.² In this market, firms are price takers and produce according to a Cobb-Douglas

²Wenzelmann and Schönfeld (2022) show that this assumption holds true for the German apprentices. In the first year of a three-year apprenticeship, for example, apprentices spend 66 percent of their working

technology:

$$y = x_{SW}^\alpha x_{AP}^{(1-\alpha)} \quad (1)$$

where $\alpha \in (0,1)$ and denotes the constant share parameter. Firms minimize their expected costs given output y ,

$$\text{minimize } w_{SW}x_{SW} + w_{AP}x_{AP}, \quad (2)$$

where the wages for both types of labor equal the marginal costs and are fixed such that $w_{SW} \geq w_{AP} > 0$. Assuming perfect competition, the marginal costs equal the marginal product for x_{SW} and x_{AP} and can be written as

$$w_{SW} = \alpha \left(\frac{x_{AP}}{x_{SW}} \right)^{1-\alpha} \quad \text{and} \quad (3)$$

$$w_{AP} = (1 - \alpha) \left(\frac{x_{SW}}{x_{AP}} \right)^\alpha, \quad (4)$$

which depend on the relative input of the other production factor.

Using Equations (3) and (4), we can deduce the expected effect of the negative supply shock of the number of skilled workers owing to emigration on skilled workers and apprentice wages. Letting $t = 0$ mark the point in time before and $t = 1$ mark the point in time after the negative supply shock, we assume that $x_{SW}^1 < x_{SW}^0$ after an outflow of skilled workers. From Equations (3) and (4), it then follows that the downward shift in x_{SW} increases the marginal product of skilled workers $\left(\frac{\partial w_{SW}}{\partial x_{SW}} < 0 \right)$ and reduces the marginal product of apprentices $\left(\frac{\partial w_{AP}}{\partial x_{SW}} > 0 \right)$. Therefore, we expect the wages of skilled workers to increase and those of apprentices to decrease: $w_{SW}^1 > w_{SW}^0$ and $w_{AP}^1 < w_{AP}^0$.

The change in the number of apprentices x_{AP} depends on adjustments to the demand and supply sides. Concerning the latter, the supply of apprentices may decrease because apprentice wages decrease, or it can increase because apprentices can expect higher wages after graduation owing to higher wages for skilled workers in the neigh-

days performing productive tasks at their workplace and 24 percent of their working days performing tasks at the skilled worker level. By the third year, the time spent on skilled worker tasks amounts to 50 percent of the total working days.

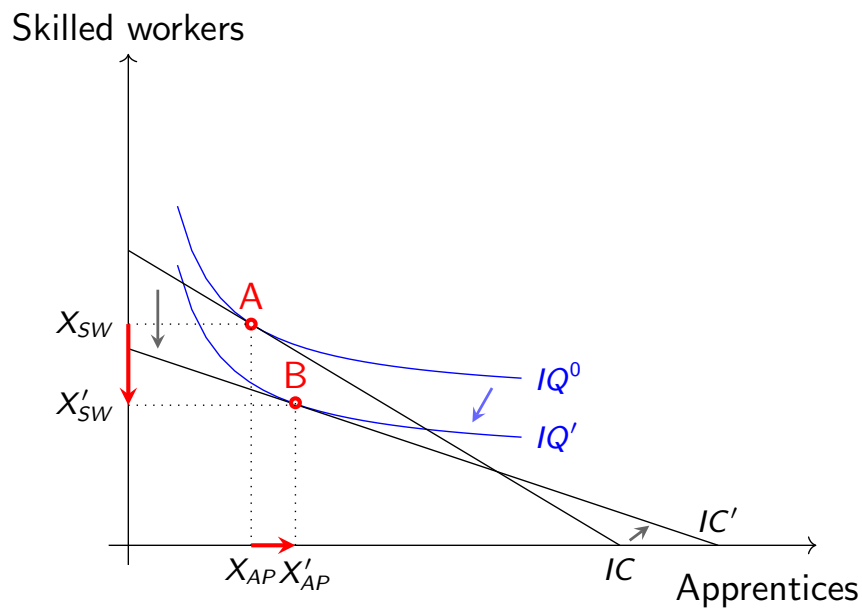


Figure 1: Adjustment of isocost curve and isoquant when apprentices and skilled workers are substitutable.

boring country. Maximizing lifetime earnings, their supply depends not only on the apprentice wage w_{AP} but also on their future wage possibilities w_{WS} . If the increase in wages by opening the borders for skilled workers is large, we thus expect that apprentice supply will increase in response to the increased future wage prospects, leading to higher apprenticeship numbers despite lower apprentice wage levels.

On the demand side, it becomes relatively costlier to employ skilled workers than apprentices because apprentice wages decline relative to skilled worker wages, thereby changing the slope of the isocost curve. To show this, we illustrate an exemplifying ex-ante equilibrium at point A, indicating skilled worker input on the y-axis and apprentice input on the x-axis in Figure 1. If, after a decrease in the number of skilled workers, their wages increase relative to the apprentice wages, the isocost curve flattens further. Firms then minimize costs at point B, leading to an input combination with a higher number of apprentices and fewer skilled workers.

In summary, we derive the following hypotheses on training in firms in the presence of a negative labor supply shock:

(H1) We expect apprentice wages to decrease.

$$H1 : w_{AP}^1 < w_{AP}^0$$

(H2) We expect the number of apprentices to increase.

$$H2 : x_{AP}^1 > x_{AP}^0.$$

III. Skilled worker outflows from German border regions into Switzerland - institutional context

In the following analyses, we test these hypotheses by considering worker outflows from the German border region to Switzerland, which provides an ideal setting to test the hypotheses derived from the textbook theoretical model. In particular, we study the introduction of the Agreement on the Free Movement of Persons (AFMP) between Switzerland and the EU, which in the early 2000s substantially facilitated cross-border commuting for so-called cross-border workers (CBW) residing in Germany and working in Switzerland. In two steps, the AFMP removed restrictions for CBWs, resulting in large outflows of German workers seeking employment in Switzerland, mainly from regions near the Swiss–German border, where commuting costs are low.

The following sections describe how we use the introduction of the AFMP to measure a causal effect of the negative labor supply shock on training in firms in the German border region. For this, we first exploit the AFMP’s stepwise introduction in different reform phases, yielding different treatment intensities over time. Second, we leverage differences in treatment intensity by travel distance by exploiting the fact that access to the Swiss labor market differs by geographic location owing to bridges, mountains, or road availability, because travel distances determine whether commuting is actually feasible in a sufficiently short time.

1. Two reform phases with varying treatment intensities

The first feature of the AFMP that we exploit is the variation in treatment intensity over time, which originates from the stepwise introduction of the AFMP that lifted commuting restrictions in two steps. Before the opening of the border (i.e., before 2002), becoming a CBW was limited and subject to specified conditions that made access to the Swiss labor market costly. The most important restriction was that companies had to comply with when they wanted to hire foreign workers, namely, the so-called *priority requirements*. These priority requirements obliged Swiss firms to prove that no suitable Swiss candidate was available for a particular job before hiring a foreign worker. A second important restriction was that cantonal authorities were required to inspect the salaries and working conditions of CBWs before issuing work permits. Moreover, CBWs were required to reside in Germany's border region, a defined number of districts right at the border (see Figure 4)³, and to commute daily to their municipality of residence. These restrictions generated substantial administrative costs for both CBWs and Swiss firms seeking to hire German CBWs.

The AFMP was signed in 1999, came into force in 2002, and gradually lifted restrictions in two phases according to Beerli et al. (2021): a Transition Phase and a Free Movement Phase. To capture the AFMP's effect on the labor market, Beerli et al. (2021) define the first phase of the AFMP as (1) the **Transition Phase** between 1999 and 2003. It starts with the announcement of the AFMP and covers anticipatory effects as well as first relaxation effects. Anticipatory effects occurred between 1999 and 2001, that is before the reform came into force. First relaxation effects occurred between 2002 and 2003 when the requirement of a mandatory daily commute was substituted with a weekly commute. Beerli et al. (2021) define the second phase from 2004 onwards as (2) the **Free Movement Phase**, which eliminated both the priority requirement granted to Swiss workers and the cantonal authorities' inspection of salary and working condi-

³The border region comprises the city of Freiburg, the city of Kempten (Allgäu), the districts Breisgau – Hochschwarzwald, Lörrach, Waldshut-Tiengen, Schwarzwald – Baar-Kreis, Tuttlingen, Konstanz, Sigmaringen, Biberach, Ravensburg, Bodenseekreis, Lindau (Bodensee), and Oberallgäu. These districts were defined by the "Abkommen zwischen dem Schweizerischen Bundesrat und der Regierung der Bundesrepublik Deutschland über den Grenzübertritt von Personen im kleinen Grenzverkehr (1970) AS 1970 1020 (CH)".

tions. This significantly reduced Swiss firms' costs of recruiting CBWs.

2. Travel distances to the border and treatment intensity

In addition to the differences in treatment intensity over time, we exploit the fact that the AFMP mainly affects firms located close to the border. Initially, only commute restrictions were lifted for residents of the border region, thereby clearly defining a treatment group for reform. However, within this border region, there are also strong differences between municipalities in the number of CBW residents and, therefore, in the actual local negative labor supply shock. These differences in the border region arise because of higher or lower travel times within a region to the Swiss labor market because of the distance to the border as well as preexisting roads and bridges across the river Rhine that were not relevant for workers commuting before the AFMP.

We use this fact to distinguish between regions of low and high treatment intensity, where assignment to groups can be seen as exogenous, assuming that access to roads and bridges across the border was irrelevant for workers' labor market decisions prior to the opening of the border. According to the Microcensus (2010), more than 70 percent of German workers commute for a time of less than 30 minutes to their workplace. However, the travel distance, respectively the commuting time from municipalities in the AFMP-defined border region to the next Swiss city varies between 5 and 105 minutes. Therefore, the AFMP provides the opportunity to commute to Switzerland within the typical 30 minutes only in some border regions. In other regions, that are municipalities located farther away, workers would have to commute, on average, longer to Switzerland than they had commuted before and longer than they typically would if they stayed in Germany. Hence, the former group can switch to Switzerland without any changes in commuting costs, whereas the latter group has higher commuting costs.

Similar to Beerli et al. (2021), we leverage this fact by using travel distance to distinguish between municipalities with high and low direct exposure to worker outflows following the AFMP. For this, we calculate travel distances in minutes by car from the

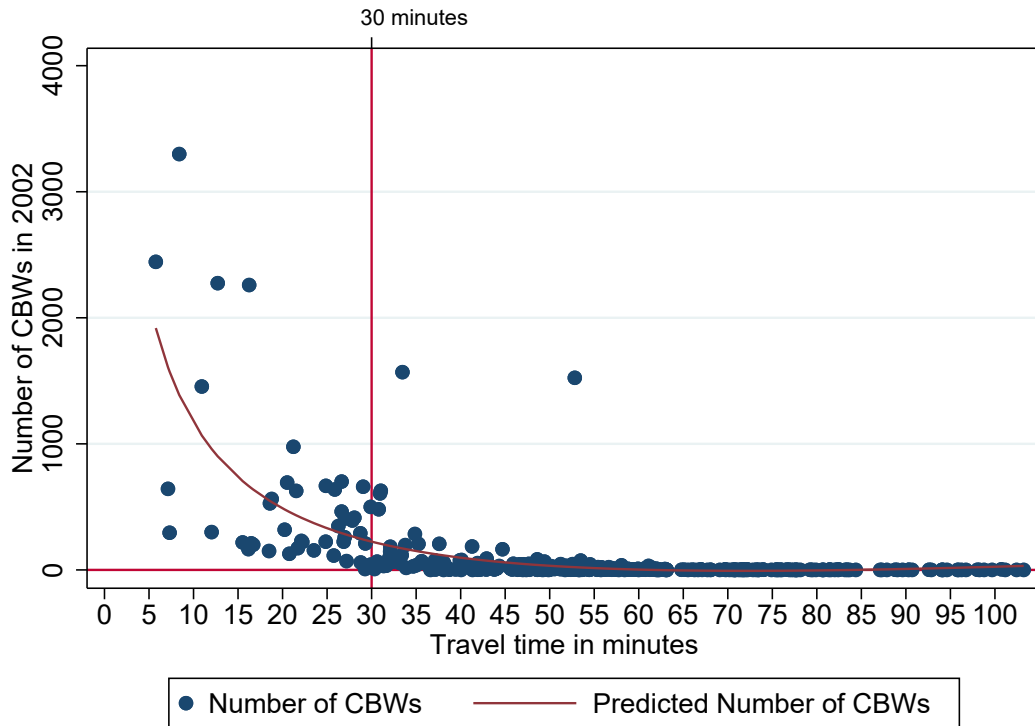


Figure 2: Districts' distance from the border and number of CBWs in 2002. Source: CCS 2002.

center of each border region municipality to the next Swiss city. The travel distances describe the time spent commuting by car and not areal distance.⁴ In this way, we could consider the actual accessibility of the Swiss labor market, which is not only determined by distance but also by transport infrastructure, such as roads and bridges.

We use the 30-minute mark as travel time to distinguish between high and low treatment intensities according to geographic location within the border region.⁵ Additionally, we find that 30 minutes of travel time is a good cut-off point, as the number of CBWs in municipalities located farther than 30 minutes away from the next Swiss city becomes very small (except for two outliers). This is shown in Figure 2, which plots the average travel distance in minutes to the next Swiss city against the number of CBWs for each municipality in the border region in 2002 and the predicted numbers of CBWs from an estimation of a fractional polynomial of travel time (red line). The

⁴We use stata's georoute package to compute travel distances which is based on HERE information.

⁵To ensure that our results were not driven by the selection of the 30-minute cut-off, we conduct additional robustness test using other cut-offs and a continuous measure of distance to the next Swiss city; see Section V.2., Figure A.2 to Figure A.4, and Table A.4.

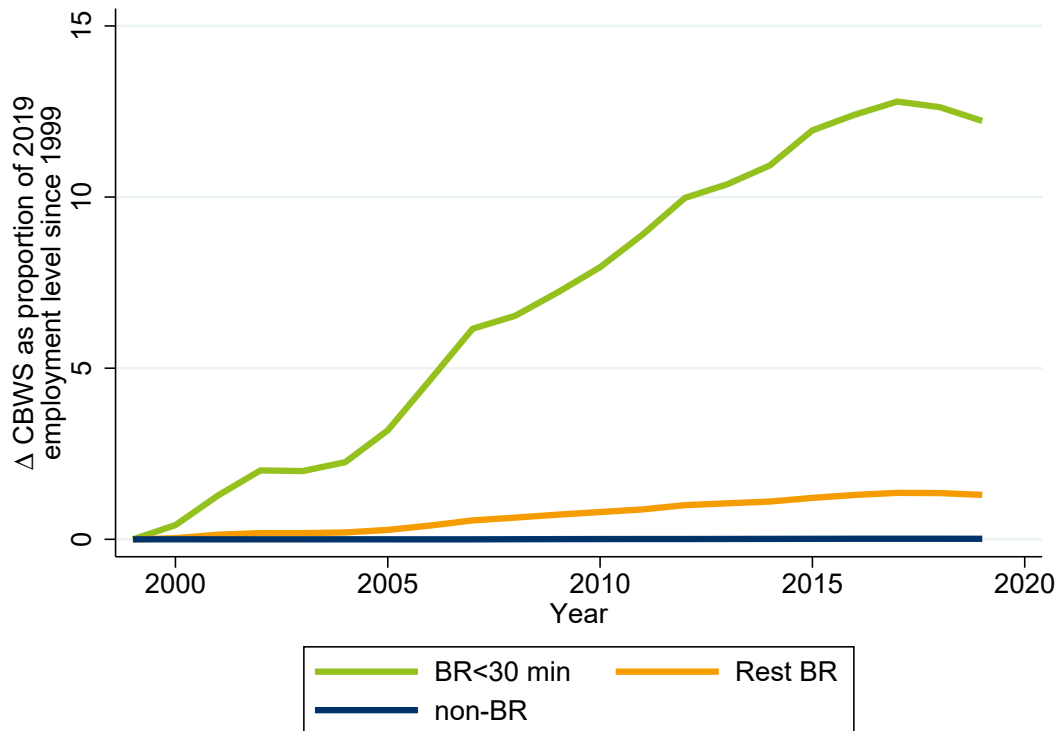


Figure 3: **Increase in CBWs in two treatment and control regions** (increase in share of migrants in percentage points relative to the base level in 1999, BR<30 min: 11.3%; Rest of BR: 0.9%; non-BR: 0.0%). Source: CCS and BeH 1999-2019.

graph shows that when the AFMP came into force in 2002, CBWs were concentrated in municipalities within a travel distance of 30 minutes to the next Swiss city (vertical line). Six of the eight municipalities with 1,000 or more CBW residents lie within this travel radius. The outliers with travel times above 30 minutes but with high numbers of CBWs were the cities of Freiburg and Waldshut-Tiengen. Past the mark of 30 minutes of travel time, the number of CBW residents in municipalities approaches zero on average. This concentration of CBWs within the travel radius of 30 minutes continued to increase after 2002. Figure 3 illustrates the increase in the share of CBW residents in the workforce from 1999 to 2019. Over this time span, the share of CBWs more than doubled in municipalities within the 30-minute travel radius, from 11.3 percent in 1999 to 23.5 percent in 2019. In contrast, border region municipalities located farther than 30 minutes away from Swiss cities experienced a very moderate development in the share of CBWs, while, outside the border in selected control regions, the share of CBWs working in Switzerland remained almost zero throughout the years.

Our selection yielded two treatment regions: (1) municipalities of the border region with a maximum travel distance of 30 minutes to the next Swiss city and (2) municipalities of the border region located farther than 30 minutes away from Swiss cities. Our control group consisted of selected municipalities located outside the border region (see Section IV.3.). Figure 4 maps the control regions and the two different treatment regions.

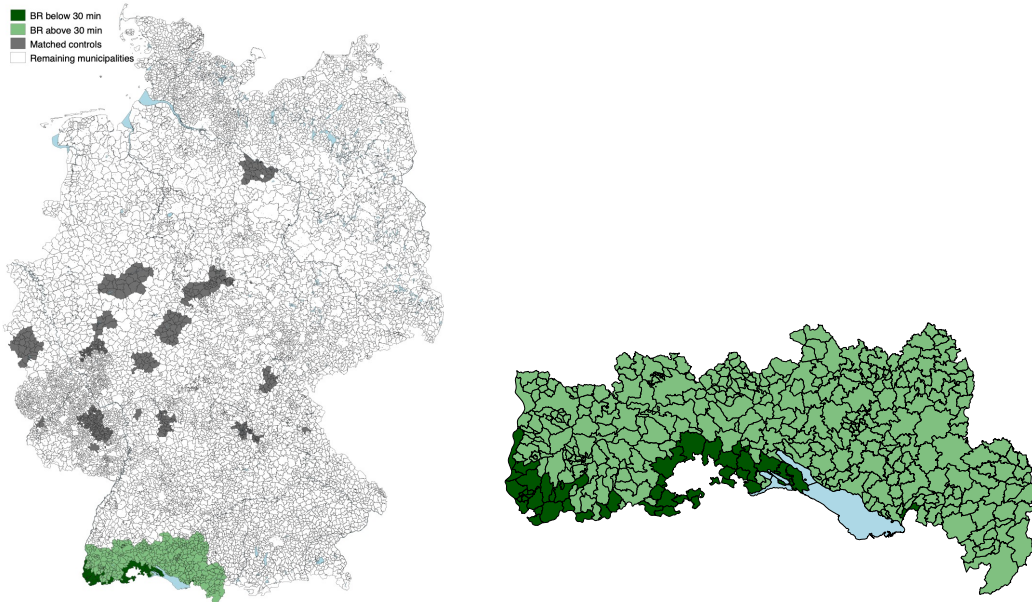


Figure 4: Treatment and control regions. Treated municipalities in the border region (green), matched control municipalities (grey), other municipalities excluded from analysis (white).

3. Benefits of cross-border work

After changing the legal framework, the AFMP induced an outflow of workers from Germany to Switzerland because of relatively high financial benefits owing to higher gross wages. The net returns are substantially higher despite the fact that the German side of the border is comparably prosperous and providing good employment opportunities. The German border region is characterized by high GDP per capita, low unemployment rates, and high average income (cf. BMI, 2020). Nonetheless, wage data show that becoming a CBW in Switzerland is attractive, especially for skilled workers. Average wages in Switzerland are roughly 36 percent higher than those in Germany (OECD, 2021). In some typical apprenticeship occupations, the wage differential is

larger. For example, between 1998 and 2002, for technicians and associate professionals, wages were, on average, 41 percent higher on the Swiss side of the border (see Table A.1). Additionally, while earning Swiss wages, CBWs in Germany benefited from lower living costs. For example, per OECD (2023), in 2021, the cost of living in Germany would be 39 percent lower than that in Switzerland. These high financial incentives to leave Switzerland as a CBW are well-known and frequently taken advantage of.⁶ In addition, searching for a job in Switzerland is comparatively easy because of the high comparability between the German and Swiss apprenticeship system. German workers also have access to information on cross-border work and support for their job search in Switzerland via several channels, such as information sessions and job boards in vocational schools, job fairs, a plethora of private agencies, and even public programs. An example of the latter is EURES (European Employment Services), a cooperation network of public employment services in the European Union (EU) and the countries in the European Free Trade Association (EFTA) aiming at facilitating the mobility of workers within these countries. EURES is active in the border region, especially in the regions of Oberrhein and Bodensee.

Given these circumstances, it becomes obvious why the number of CBWs increased so dramatically after the announcement of the AFMP in the typical occupational fields of apprenticeship graduates (see Figure A.1). Hence, the AFMP provided highly profitable opportunities for workers with completed apprenticeships in the Swiss labor market. This, in turn, also affected the benefits of choosing an apprenticeship for young adolescents because it substantially improved their future labor market prospects.

4. AFMP as a natural experiment

In summary, we argue that the AFMP is an excellent setting for estimating the effects of a negative labor supply shock on training in firms for two reasons. First, the resulting negative labor supply shock in Germany provides a clear exogenous variation in the local skill supply, allowing us to identify the causal effects. We were not only

⁶As represented by a recent news article in Vuillemin (2023) titled "Since I work in Switzerland, I wallow in luxury."

able to leverage differences in treatment over time but also intensity by distance from the border. Second, the AFMP provided incentives for young adolescents to train for becoming skilled workers because of the high financial incentives when working as a CBW after completing their apprenticeship. Therefore, the reform is likely to directly affect the formation of human capital for skilled workers through apprenticeship training in the sending country.

IV. Empirical strategy and data

In this section, 1. we outline our empirical specification based on the institutional context of the reform, considering the two sources of variation in treatment intensity (over time and by geolocation of firms) using a DiD approach. 2. We introduce the database used in the estimation and its preparation to suit the empirical approach. 3. We explain our control group selection (see Figure 4), which we use as a point of reference to estimate and interpret the effects of the AFMP. Finally, we discuss the matching quality.

1. Empirical specification

To estimate the effect of the AFMP on the number of apprentices and apprentice wages in firms, we apply a DiD framework that considers the differences in treatment intensity according to the two reform phases of the AFMP and the location of firms in terms of travel distance to the border. As in any standard DiD design, we are interested in the difference in our outcomes of interest in the treated regions compared with the control regions before and after the reform. In our case, we not only evaluate differences before 1999 and after the reform but also distinguish between two phases after the reform: (1) The dummy variable *Transition* indicates whether observation time t concerns the years 1999 to 2003 and (2) the dummy variable *Free* indicates that the observation is made from 2004 to 2019.

We then distinguish between two different treatment groups: (1) firms located in

the border region and within 30 minutes of travel distance to the next Swiss city, as indicated by the dummy variable $\mathbb{1}[BR, d_{i(m)} < 30]$ and (2) firms located within the border region but farther than 30 minutes away, as indicated by the dummy variable $\mathbb{1}[BR, d_{i(m)} > 30]$, where distance d of firm i depends on their location given municipality m . In contrast, the control firms are located in selected regions that have similar labor market characteristics to the treatment regions but are located outside the border region (Section IV.3. describes this matching approach in greater detail).

As in standard DiD, we are interested in the average treatment effect, which is measured by the interaction coefficients of the treatment and post-reform indicators. In our model, these are signified by γ_1 to γ_4 , yielding the average treatment effects in the Transition Phase for a firm located within 30 minutes of travel distance (γ_1) and for firms located within the border region but farther away (γ_2), as well as the effect in the Free Movement Phase for firms located within 30 minutes of travel distance (γ_3) and for firms located farther away (γ_4). Therefore, our model takes the following form:

$$\begin{aligned}
Y_{it} = & \gamma_1 Transition_t \times \mathbb{1}[BR, d_{i(m)} < 30] + \gamma_2 Transition_t \times \mathbb{1}[BR, d_{i(m)} \geq 30] \quad (5) \\
& + \gamma_3 Free_t \times \mathbb{1}[BR, d_{i(m)} < 30] + \gamma_4 Free_t \times \mathbb{1}[BR, d_{i(m)} \geq 30] \\
& + \alpha_t + \alpha_i + \beta X_{i(m)t} + \epsilon_{it},
\end{aligned}$$

where we control for year fixed effects (α_t) and firm fixed effects (α_i), as well as linear time trends in the broader NUTS-II region ($X_{i(m)t}$), where the firm captures general time-varying but location-specific trends. Overall, our model specification is similar to that proposed by Beerli et al. (2021). We cluster the standard errors at the level of 61 municipal associations. As our clusters are unbalanced in size, we report the p-values obtained from a wild bootstrap procedure (Roodman et al., 2019) with 501 repetitions.

2. Data

Certain data requirements must be satisfied to estimate Equation 5. First, we need firm-level data on apprenticeship training that cover not only the entire period of the two

reform phases but also sufficient observation time before and after the announcement of the AFMP. Second, we need data that provide detailed information on firm location and a sufficiently large sample of firms in comparably small border regions.

Therefore, for our analyses, we use administrative data from the Institute for Employment Research (IAB)⁷ for 1990 to 2019. We focus on Employee History (Beschäftigtenhistorik - BeH), which contains the annual employment and end-of-employment notifications of all employees covered by social security (since 1975) and employees in marginal part-time employment (since 1999) submitted by employers to responsible social security agencies. This includes the vast majority of workers in Germany and their apprentices.⁸ The data can be combined to construct labor market biographies of workers as well as employment structures and turnover at the establishment level.

These data offer several advantages to our analysis. (1) Firms can be located at the municipality level, making it possible to assess their distance from the next Swiss city and thereby identify how strongly they were effected by the AFMP in what year. (2) BeH is the largest available database on employment with this level of regional detail. We were able to use a 20 percent sample that ensures a sufficiently high representation of the border region over a long period. (3) The data provide reliable information on the number of employees and apprentices as well as their gross daily wages. This enables us to analyze both employment and wage effects for apprentices.

We use a 20 percent sample of the entire universe of establishments that appeared in the registries on June 30, 1998, and are located either in the border region or in a matched control region. For these firms, we observe all employees who were employed at the firm at some point between 1990 and 2019.⁹ Using these data, we construct a panel of establishment observations on the reference date September 30 in each year between 1990 and 2019 while only considering firms that did not move into or out

⁷The data are first registered by the health insurance companies, collected and processed by the German Federal Employment Agency (BA - Bundesagentur für Arbeit), and then integrated into the Employee History File by the IAB (Vom Berge et al., 2021).

⁸Self-employed persons, civil servants, and unemployed students are not included in the data because they are exempt from social security contributions.

⁹We observe the labor market biographies of all employees after the year 1990 and after they started to work in this firm; that is, we observe job changes in other firms until 2019 but not the employment history before starting to work at the sample firm.

of their municipality during the entire observation period.¹⁰ We focus our analysis on private-sector firms and exclude the public administration sector because its labor market is too particular to be comparable to the rest of the economy.¹¹ Furthermore, we conduct our main analyses excluding the city of Freiburg, which presents itself as an outlier in three dimensions. 1. Freiburg's industrial structure is very different from that of the rest of the border region: It has a very strong health and social sector, which would consequently dominate our sample observations.¹² 2. Freiburg hosts a large share of border region firms, almost as large as that of border regions within the 30-minute travel band. 3. Freiburg is located farther than 45 minutes away from the next Swiss city despite large numbers of CBW residents, revealing that residents of Freiburg are willing to take on longer travel times than average. To ensure that our results are not dominated by developments exclusively attributable to Freiburg and that there is overall a high consistency across the two border region zones, we exclude the city from the sample. However, we show results including Freiburg in Table A.5.

At the reference date, we observe the number of apprentices employed at the firm and their gross daily wages, which we further process in the following ways to meet our data requirements: 1. Considering that an apprenticeship takes between two to three-and-a-half years to complete, we exclude apprentices with a training duration of more than five years. However, this restriction excludes only approximately 1.1 percent of the sample's apprenticeship observations. 2. Because most firms that participate in apprenticeship training are small (cf. Table 1) and do not hire new apprentices on a yearly basis, we smooth the number of apprenticeships as a three-year moving average, resembling the average program duration of an apprenticeship. We proceed in a similar manner to apprentice wages and apprentice retention rate.¹³ In general, we

¹⁰We do not restrict the analysis to a balanced sample of establishments. This means that we include establishments that close-down during the observation period.

¹¹The adjustment of apprenticeship demand in the public administration sector is likely to show a lower correlation with changes in the economic environment. Wenzelmann and Schönfeld (2022), for example, show that the public sector, on average, incurs much higher net costs per apprentice. Therefore, we exclude this sector from the analysis to avoid deluding the estimated effects.

¹²The health and social sector accounted for 49 percent of employment in the sample firms in Freiburg in 1998, while it was only approximately 18.4 percent in the rest of the border region.

¹³Observations of zero apprentices and retained apprentices (but not for wages) are changed to a very small number in order not to lose the observation when taking logarithms. This is necessary because

deflate wages and impute values where wages are censored at the social security limit, using the standard technique proposed by Gartner (2005). Overall, we closely follow Dauth and Eppelsheimer (2020) in our data processing.

We observe data from more than 20,000 firms between 1990 and 2019, of which approximately 12,000 lie within the legally defined border region and almost 2,200 within a 30-minute travel radius to the next Swiss city. Within these firms, the data comprise information on about 171,000 apprentices, of which approximately 98,000 live in the border region and approximately 15,500 within a 30-minute travel radius to the next Swiss city.

3. Matching control regions

To evaluate the treatment effects in the border regions, we must compare their development with those in comparable untreated regions. For this, we do not use all available control regions but apply nearest-neighbor matching, which ensures data sparseness by limiting the number of control regions and the pre-reform comparability of treatment and control regions by selecting control regions that are as similar as possible to the border region in terms of labor market characteristics.

To select good matches, we use Mahalanobis distance matching with replacement, where we match the nearest neighbors with similar features for the border region districts. We use district-level data for 1998, the year before the announcement of the AFMP. Using the IAB Establishment History Panel (BHP)¹⁴ aggregated at the district level, we match the districts' total employment levels, shares of low and high skilled workers, shares of employment in manufacturing, logarithms of mean worker wages, and proportions of right-censored wages.¹⁵ We also account for population density, net numbers of cross-district commuters, and levels of apprenticeship employment in

some firms only train sporadically or decide to participate in training only later in the observation period. Since our dependent variable is the logarithmic number of apprentices, these observations would be lost.

¹⁴The data include the entire universe of German establishment in 1998.

¹⁵Administrative records only document wages as long as they are subject to social security contributions. Higher-income earners are eligible for private social security. Censored wages above this social security limit were imputed in the data set.

1998, which we retrieve from the INKAR database.¹⁶

To ensure that the control regions are unaffected by the AFMP (e.a., by spill-over effects from neighboring regions), we only consider matches located at least three hours away from the Swiss border. However, to select districts that are as similar as possible, we only consider other districts in West Germany and not in East Germany because of the persistent underlying structural differences in apprentice demand and supply between West and East Germany (see for example Riphahn and Zibrowius, 2016). For each border region district, we select as many nearest neighbors as possible outside the border region until the sum of their employment levels equals at least 1.5 times that of the respective border region districts, in line with Dustmann et al. (2017). In total, we match 19 control districts comprising 555 municipalities.

Overall, the resulting matched control regions were comparable with the border regions. Table A.2 shows that, for 1998 (the year before the announcement of the AFMP), by considering the selected control regions instead of all German districts, we could compare districts with a higher degree of similarity, especially with regard to population density, but also characteristics that were unaccounted for, such as the unemployment rate, average number of firms in a district, and share of manufacturing employment.

Table 1 shows that our approach yields a good match when examining the firm-level sample used in our analyses. The table provides descriptive statistics for apprentices and workers who worked in the border region with less than 30 minutes and more than 30 minutes of travel time to the next Swiss city, respectively, compared with the matched control municipalities for the year 1998. Evidently, the control regions and two treatment regions are similar with respect to apprentice characteristics, the share of manufacturing apprenticeships, and mean apprentice wages. Notably, apprentices in the border region are employed slightly more often in manufacturing and smaller es-

¹⁶INKAR is operated by the Federal Institute for Research on Building, Urban Affairs and Spatial Development. The population density is based on employment data from the Federal Employment Agency and the land area survey of the Federal Statistical Office and the Statistical Offices of the Länder. The data on apprenticeship employment and net commuters are based on surveys conducted by the Federal Employment Agency.

Table 1: Worker and Apprentice characteristics in 1998 by distance to Switzerland.

	$\mathbb{1}(BR, d < 30)$		$\mathbb{1}(BR, d \geq 30)$		non-BR	
Apprentices						
Share Abitur	0.12	(0.33)	0.11	(0.31)	0.16	(0.36)
Share foreign	0.10	(0.30)	0.08	(0.27)	0.06	(0.23)
Share female	0.48	(0.50)	0.44	(0.50)	0.42	(0.49)
Share est. Size <50	0.60	(0.49)	0.58	(0.49)	0.54	(0.50)
Share est. Size >50	0.40	(0.49)	0.42	(0.49)	0.46	(0.50)
Share manufacturing	0.27	(0.44)	0.28	(0.45)	0.24	(0.42)
Mean log. apprentice pay	3.25	(0.29)	3.25	(0.30)	3.20	(0.32)
Observations	1,384		6,358		7,042	
Workers						
Share skilled workers	0.76	(0.43)	0.77	(0.42)	0.76	(0.43)
Mean log. wages of skilled workers	4.53	(0.41)	4.55	(0.41)	4.53	(0.41)
Observations	18,748		82,399		100,707	
Firms						
Share training firms	0.24	(0.43)	0.27	(0.44)	0.27	(0.44)
Mean number of apprentices						
All firms	0.66	(3.18)	0.77	(3.34)	0.90	(5.96)
Training firms	2.72	(5.99)	2.88	(5.97)	3.30	(11.09)
Observations	2,082		8,244		7,847	

Note: Described sample restrictions apply. Source: BeH 1998.

tablishments than in matched control municipalities. However, with small differences, the selection of the control region is comparable to that of the border region.

V. Results

In this section, we first test our two hypotheses and present the estimated effects of the AFMP on apprenticeship employment and wages. Second, we present the results of several robustness checks to check whether our results are sensitive to our different operationalization or sampling decisions. Third, we provide results from further analyses of the underlying mechanisms and finally discuss the remaining general limitations of our data analysis.

1. Effect of the AFMP on training

With respect to our hypotheses derived in Section II., we expect that the introduction of the AFMP and the resulting negative labor supply shock has two measurable effects: (H1) The apprentice wage decreases in line with their marginal productivity as they become relatively more available than skilled workers. (H2) The number of apprentices increases when they substitute skilled workers with a decreasing relative cost of employment.

Given the institutional context of the AFMP, we expect that the effects may have been stronger after 2004 in the Free Movement Phase because more commuting restrictions were lifted in 2004. We also expect that the AFMP would have stronger negative effects on apprentice wages and stronger positive effects on the number of apprentices in firms located within 30 minutes of travel distance to the next Swiss city, which experienced a higher increase in the number of CBWs after the reform.

We test the two hypotheses and our expectations for treatment intensity using Equation 5. Table 2 presents the estimated treatment effects in the Transition (top panel) and the Free Movement Phase (bottom panel), for firms located in border regions with less than 30 minutes of travel time to the next Swiss city and for those located farther away, respectively.

In line with the first hypothesis (H1), our results confirm a small negative effect of the AFMP on apprentice wages. Column 1 of Table 2 indicates a decrease in apprentice wages of about 2 percent in the Free Movement Phase in firms located within 30 minutes of the next Swiss city and about 0.5 percent in border region firms located farther away. This result also confirms our expectations that the effect of the AFMP on apprentice wages was stronger when more restrictions were lifted in the second reform phase and also stronger in firms located closer to the border. Considering the average monthly apprentice wage of about 780 euros before the reform (see Table 1)¹⁷, an effect of -2 percent translates into a decrease of 15.60 euros per month (as estimated for firms close to the border in the Free Movement Phase); a wage decrease of 0.5 percent (as

¹⁷Mean gross daily wages in the border region sample approximately equal 26 euros.

Table 2: Effect of the AFMP on training in firms.

	(1) Apprentice wage	(2) Number of apprentices	(3) Probability of training
$Transition \times \mathbb{1}(BR, d < 30)$	-0.009 ⁺ (0.005) [0.086]	0.170 (0.108) [0.126]	0.011 ⁺ (0.008) [0.098]
$Transition \times \mathbb{1}(BR, d \geq 30)$	-0.007 (0.007) [0.201]	-0.018 (0.099) [0.878]	-0.003 (0.007) [0.705]
$Free \times \mathbb{1}(BR, d < 30)$	-0.020* (0.008) [0.024]	0.458** (0.165) [0.008]	0.028** (0.010) [0.006]
$Free \times \mathbb{1}(BR, d \geq 30)$	-0.005* (0.007) [0.391]	0.368** (0.126) [0.012]	0.021** (0.008) [0.016]
Observations	130,335	332,138	332,138
R2	0.760	0.684	0.560

Notes: Estimation following specification of equation (1), dependent variables: (1) logarithm of mean apprentice wages, (2) logarithm of the mean-averaged number of apprentices, and (3) training in firm (Dummy-variable equal to 1 if firm trains apprentices), controls for firm fixed effects, year effects, and a NUTSII-regional time trend, excluding the city of Freiburg. Robust standard errors in parentheses adjusted for 61 municipal association clusters, wild bootstrap p-values in brackets, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Source: BeH 1990-2019.

estimated for firms farther away) would translate to a decrease in monthly apprentice wages of approximately 3.90 euros a month.

Furthermore, the results of the second hypothesis (H2) confirm the positive effect of the number of apprentices in firms. In Column 2 of Table 2, we find evidence of a positive effect of the AFMP on the number of apprentices employed by a firm in both parts of the border region in the Free Movement Phase. On the intensive margin, apprenticeship training increased by approximately 45.8 percent in firms located very

close to the border and by 36.8 percent in firms located in the border region but farther than 30 minutes away. The effect of the AFMP on the number of apprentices, in fact, reveals that the positive effect (H2) is, as expected, stronger during the Free Movement Phase. The results support our expectation that the effect was stronger for firms close to the border. Given that training firms in the border regions employed on average two to three apprentices in 1998, the result means that firms employ about one additional apprentice after the border opening, closer to the border a bit more and farther away from the border a bit less than one on average (see Table 1).

We also find the expected positive effect on the extensive margin in Column 3, indicating that the number of firms employing at least one apprentice increased. In the Free Movement Phase, the probability of training increased by 2.8 percentage points for firms located close to the border and by approximately 2.1 percentage points for firms located farther away from Swiss cities.¹⁸ With a training share of approximately 24 percent in the border regions closer than 30 minutes travel time to the border this result indicates an increase of 12 percent in the number of firms that are willing to train apprentices; in border regions located farther away, the training share among firms was 27 percent in 1998 suggesting an increase of about 8 percent in the number of training firms (see Table 1).

2. Robustness checks

We conduct several tests to evaluate the robustness of the results. First, we test the extent to which our results are sensitive to different categorizations of the two zones of the border region. We repeat the analyses using different cut-off travel times between 15 and 45 minutes to categorize firms into two different zones of the border region (see Figures A.2–A.4). The negative effect on apprentice wages in the Free Movement Phase for firms located closer to the border is more or less insensitive to the choice of the cut-off travel time. However, the effects on the number of apprentices and training probability are stronger when the cut-off travel time is smaller. Table A.4 presents the

¹⁸We also estimate the coefficients in Column 3 using a logit model, which yields similar conclusions (see Table A.3).

results of the estimations using a continuous measure of travel time from the firm's location by municipality. Again, the results support our findings that a firm located farther away from the border is associated with significantly higher apprentice wages, lower apprentice numbers in firms, and a lower training probability in the Free Movement Phase. Thus, although our effect would be stronger if we chose a shorter cut-off time, we chose the 30 minutes travel time for our main analysis to not inflate our results and represent typical travel times.

Second, we investigate whether our results are driven by the two outliers, Freiburg and Waldshut-Tiengen, both of which are located farther away than 30 minutes but show a very particular industry structures and commuting patterns and relatively high numbers of CBW (cf. Figure 2). First, we repeat our analyses by including the city of Freiburg. Our results are largely confirmed, as we find again the positive effect on the intensive and extensive margins of training in the border region and a small negative effect on apprentice wages, which is statistically significant at the 10 percent level. In addition to Freiburg, we also test the robustness of our results by excluding Waldshut-Tiengen, the other large municipality with high numbers of CBW from the second zone of the border region. Again, our results remain robust, as shown in Table A.6.

Overall, our tests show that the main results are robust to different specifications and samples.

3. Further analyses

We conduct further analyses to shed light on the mechanisms leading to the observed effects.

First, we present some indications that the positive effect on the number of apprentices is likely driven by the increased demand from firms and at the same time an increase in the supply of apprentices (adolescents applying for apprenticeship positions). Unfortunately, with our data (or any other available data) it is impossible to observe both effects separately on the regional level that we need for our analysis. We can only observe the net effect of the possible demand and supply shifts. However, we can ob-

serve changes in the retention rate, that is, the share of apprentices who leave after graduation. If after the reform adolescents become more interested in apprenticeships because of better wage prospects in Switzerland, we would expect more graduates to leave their training firms for outside jobs. Therefore, a negative effect of the reform on the retention rate indicates that the effect on the number of apprentices was partially supply-driven. We estimate the effects of the AFMP on the firm-level retention rate of apprenticeship graduates within their training firms one, two, and three years after graduating. We find that, while the number of apprentices increased as shown in our main analysis, the retention rate decreases significantly in the border region after the opening of the borders (see Table A.7). This result supports the hypothesis that apprenticeships become more attractive to young school leavers because they use these opportunities to gain higher wages on the Swiss side of the border after finishing an apprenticeship. Thus, long-term labor market prospects strongly affect the application for apprenticeships, with high-wage options for skilled workers leading to larger numbers of applicants.

Simultaneously, the lower retention rates are consistent with the observed negative effect on apprentice wages. With apprentice graduates leaving their training firms more frequently shortly after graduation, net training costs are rising because the training costs can no longer be recouped by benefits after the training phase if apprentices leave the company. As a substantial share of training firms in Germany incur large training costs during the apprenticeship (Mohrenweiser and Backes-Gellner, 2010; Wenzelmann and Schönfeld, 2022) it is to be expected that a reduction in the average retention rate leads to an on average lower apprentice pay. Firms that cannot recoup their training costs by retaining their apprentice as a skilled worker after graduation by for example saving on wages or screening costs have increased net costs after the opening of borders. To meet these rising costs, apprenticeships must become more profitable during the apprenticeship, which, however, pressures firms to lower apprentice wages. Apprentices are willing to tolerate such lower wages because of substantially increased wage prospects after their apprenticeship (and because training

quality is guaranteed by tight institutional guard rails in the German apprenticeship system).

We also check whether the negative effect of the AFMP on apprentice wages might only reflect a composition effect, meaning a change in the occupational composition towards low-wage apprenticeships. Table A.8 shows that the changes in the occupational composition of apprenticeships in the border region are small and do not coherently fit the pattern of increasing apprentice employment in low-wage occupations. Therefore, we find no evidence that the negative wage effect for apprentices results from a composition effect; instead, it is more likely caused by an actual downward adjustment of apprentice wages, which is consistent with the theory we use.

Nevertheless, one could criticize that the wage effect is too moderate.¹⁹ However, we argue that it is very plausible because there are institutional explanations for downward rigid wages in the German system. Most apprentice wages are set through collective bargaining and the agreements are binding, even for firms not covered by the agreement.²⁰ Our results align with the recent literature analyzing the training market responses to positive supply shocks of potential apprenticeship applicants; these studies find apprentice wages to react very moderately or even insignificantly (Dorner and Görlitz, 2020; Muehlemann et al., 2022). Although the wage effect is small, our results are consistent with firms substituting apprentices for skilled workers at lower wage costs, enabled by the increasing attractiveness of apprenticeships for adolescents owing to increasing wage options after graduation.

Finally, to verify the theoretical prediction that firms substitute apprentices for skilled workers to save on wage costs, we analyze the corresponding effects on skilled worker employment and wages. In line with theoretical expectations, skilled worker employment decreases in the Transition Phase in firms close to the border (see Column 1 of Table A.9). However, we do not find a significant effect on skilled worker employment in the Free Movement Phase, indicating that the main effect on skilled worker

¹⁹Beerli et al. (2021) find a positive effect of 4.5 percent on real wages for workers on the Swiss side of the border.

²⁰Uncovered firms must not undercut apprentice wages set by collective agreements in the same region and industry by more than 20 percent.

employment occurred rapidly after the announcement of the AFMP. We find a negative average effect on the wage rate of skilled workers of about -2.5 percent in firms located close to the border and of about -1.7 percent in firms located farther away in the border region (see Column 2 of Table A.9). The negative effect is driven by a change in the composition of the skilled workforce within firms, which is due to workers with the highest pay leaving to work in Switzerland while lower-paid workers stayed. This result aligns with Borjas (1987), who discusses the Roy (1951) model, showing that emigrants may be positively selected with respect to earnings if net mobility costs, that is, the expected wage increase when going abroad, are the highest for the highest-paid workers.²¹ At the same time there is no negative effect for incumbent workers who worked in the firms before the reform. If at all, our results indicate small positive effects for incumbent workers close to the border, with a positive coefficient at a significance level close to the 10 percent threshold (see Column 3 of Table A.9). Moreover, the average wage level of skilled workers is significantly higher than that of the control regions immediately after 2004, that is, at the beginning of the Free Movement Phase (see Figure A.5). Therefore, the results largely support the theoretical hypothesis that skilled worker wages increased after a negative supply shock of skilled workers in the border regions.

4. Limitations

Although further analyses shed light on the underlying mechanisms, we need to address the limitations of our study.

First, with the available data, we cannot disentangle the supply- and demand-side effects of policy reforms. The main reason is a lack of high-quality supply-side data on apprenticeship applicants. Survey-based data are not helpful because the number of observations in the border region is too small. Alternative administrative records

²¹There has been some controversy concerning the model because, in many empirical examples, emigration is positively selective, even though the returns to education are lower in the host country. That is, mobility costs are actually higher for the well-educated. More recently, Leopold et al. (2023) show that these differences in mobility costs can be explained by differences in returns on occupational skills using the example of Mexican migration to the US.

on regional apprenticeship supply and demand are available only after 2008 and after the AFMP affected apprenticeships in the border regions. Finally, the administrative data we used for analysis provide limited information on the circumstances of entering apprenticeships and exiting the German labor market. Individuals only entered our data with their first social-security-relevant job, which was an apprenticeship for the majority. Therefore, we could not analyze the determinants of apprenticeship choice. For the same reason, we could not test whether apprentices left their training firms to become CBWs in Switzerland. Moreover, high school diplomas are recorded noisily, which is why we refrained from conducting a more detailed analysis of the effects of school qualifications.

The second limitation concerns the investigation of the potential spillover effects in adjacent regions. As expected, we find stronger effects for firms located within 30 minutes of travel distance to the border. However, we also found strong effects on the apprentice numbers in border regions located farther away. This could point to cascade effects, where the AFMP also indirectly affected firms in locations farther away from the border with low actual labor outflow to Switzerland because the outflow of workers in close border regions was replaced by a new inflow of workers from regions located farther away. In other words, for residents of these regions, travel times to Switzerland may be too long; however, travel times to municipalities located close to the border may be short enough for them to seek work in firms located in these areas. Ultimately, a detailed analysis of these dynamics is beyond the scope of this study and will be left to the future.

VI. Conclusion

This study investigates how international mobility and particularly the outflow of workers affect training adjustments in firms in the sending country. Our example are middle skilled workers who are trained in apprenticeship training programs by German firms. Our analysis builds on a reform that opened the Swiss border for cross-

border commute and led to a considerable exogenous negative labor supply shock in Germany, as workers commuted in large numbers from German border regions to Switzerland, taking advantage of substantially higher wages in the neighboring country.

For the empirical analysis, we applied a DiD framework by leveraging the different dimensions of treatment intensity associated with the reform. On the one hand, we exploited the difference in treatment intensity over time owing to a step-by-step introduction of the border opening, which lifted the restrictions for cross-border commuters gradually in two steps. On the other hand, we used the variation in the treatment intensity (i.e., the size of the negative labor supply shock) by the location of the firm in terms of travel time to the next Swiss city.

We find a robust increase in the number of apprenticeships despite lower apprenticeship wages. What seems like a puzzle can be explained based on our theoretical model, suggesting that the results are due to a mix of supply- and demand-side adjustments: On the demand side, firms increase their demand for apprentices to fill the gap in skilled workers. They pay lower apprenticeship wages to compensate for lower post-training benefits and the risk of lower retention rates due to the increased brain drain.

On the supply side, potential apprenticeship applicants have higher incentives to apply for apprenticeship training openings because the expected wage and other returns after graduation increase owing to the opening of the border. By weighing lower short-run training wages against substantially higher long-run worker wages and benefits across the border, young workers are more inclined to start an apprenticeship. Our evidence for the increased post-training mobility of former apprentices after the policy reform lends some support to this interpretation.

Our study contributes to the literature in three ways. First, it contributes to the research on the effects of migration of workers between industrialized countries, especially to the low number of studies addressing the effects of training in firms in the sending country. We also complement the classic literature on labor demand and sup-

ply models by presenting a textbook application example of a model with input adjustments in tight labor markets. Finally, we contribute to training literature by explaining investments in training among individuals and firms.

Our results have implications for both policy and practice. The German training market is suffering from excess demand for well-qualified and motivated apprentices and a lack of apprenticeship applicants. Consequently, several industries lack young workers to meet the skill demands of firms. At the same time, the real wages of apprentices and skilled workers have decreased in recent years. Our results suggest that increased skilled worker wages not only incentivize skilled workers but also attract young adults to choose a vocational career as offered by apprenticeships. Thus, attractive skilled worker wages and other benefits seem to be one of the key mechanisms for attracting more applicants to the German - or any other - training market and solving the large shortages of skilled workers in the long run.

Our research findings could be generalized in two important ways. They can be expanded beyond Germany and Switzerland to other nations facing similar skilled labor dynamics. The theoretical model can be applied to many countries experiencing brain drain due to better opportunities of skilled workers abroad, impacting their domestic labor supply. The key parameters — such as the wage gap between sending and receiving countries, labor supply shocks, and post-training benefits — could be adapted for different countries and many industries in which companies are investing in workers' training in the form of apprenticeship training or similar middle skill training measures. However, our results also indicate that while brain drain can present challenges, it can also yield brain gains because it encourages investments in education and increases the supply of skilled workers in the home country.

Moreover, our research findings could be generalized to other sectors and other levels of education with similar training structures, in which part of the training takes part while working on the job (e.g. PhD candidates, auditors, lawyers or teachers). The academic labor market, for example, employs a similar apprenticeship-style training system in the form of PhD programs where doctoral candidates work for low wages

or grants to receive an education in a university while working with established researchers on their research projects. Brain drain is a well-documented phenomenon in academia as well and our results suggest that it is the combination of wages during the training phase in comparison to labor market opportunities after finishing the training phase (at home or abroad) that determine the attractiveness of PhD training for individuals and universities alike. These examples show that our findings are generalizable in many ways and, hence, go far beyond the German-Swiss border regions or the example of apprenticeship training programs.

References

- Aeppli, M. and A. Kuhn (2021). Open labor markets and firms' substitution between training apprentices and hiring workers. *Labour Economics* 70, 101979.
- An, L., Y. Qin, J. Wu, and W. You (2022). The local labor market effect of relaxing internal-migration restrictions: Evidence from China. *Journal of Labor Economics* (accepted manuscript).
- Berli, A., J. Ruffner, M. Siegenthaler, and G. Peri (2021, March). The abolition of immigration restrictions and the performance of firms and workers: Evidence from Switzerland. *American Economic Review* 111(3), 976–1012.
- BMI (2020). Deutschlandatlas 2019. Technical report, Bundesministerium des Innern, für Bau und Heimat.
- Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *American Economic Review* 77(4), 531–553.
- Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The Quarterly Journal of Economics* 118(4), 1335–1374.
- Bütikofer, A., K. V. Løken, and A. Willen (2019, Oct). Building bridges and widening gaps: Efficiency gains and equity concerns of labor market expansions. Discussion paper series in economics, Norwegian School of Economics, Department of Economics.
- Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *ILR Review* 43(2), 245–257.
- Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics* 19(1), 22 – 64.
- Dauth, W. and J. Eppelsheimer (2020). Preparing the sample of integrated labour market biographies (siab) for scientific analysis: a guide. *Journal for Labour Market Research* 54(10), 1–14.

- Dicarlo, E. (2022). How do firms adjust to negative labor supply shocks? evidence from migration outflows. IZA Discussion Papers 14994, Institute of Labor Economics (IZA).
- Dorner, M. and K. Görlitz (2020). Training, wages and a missing school graduation cohort. *IZA Discussion Paper* (13490).
- Dorner, M. and K. Görlitz (2020). Training, wages and a missing school graduation cohort. IZA Discussion Papers 13490, Institute of Labor Economics (IZA).
- Dustmann, C., F. Fabbri, and I. Preston (2005). The impact of immigration on the british labour market. *Economic Journal* 115(507), F324–F341.
- Dustmann, C. and A. Glitz (2015). How do industries and firms respond to changes in local labor supply? *Journal of Labor Economics* 33(3), 711–750.
- Dustmann, C., U. Schönberg, and J. Stuhler (2016, November). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives* 30(4), 31–56.
- Dustmann, C., U. Schönberg, and J. Stuhler (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics* 132(1), 435–483.
- Gartner, H. (2005). The imputation of wages above the contribution limit with the german iab employment sample. FDZ-Methodenreport 02/2005, Institute for Employment Research IAB.
- Glitz, A. (2012). The labor market impact of immigration: A quasi-experiment exploiting immigrant location rules in Germany. *Journal of Labor Economics* 30(1), 175 – 213.
- Hafner, F. (2022). The equilibrium effects of workers' outside employment options: Evidence from a labor market integration. Working paper.
- Leopold, S., J. Ruhose, and S. Wiederhold (2023). Why is the Roy-Borjas model unable to predict international migrant selection on education? Evidence from urban and rural Mexico. *IZA Discussion Paper* (16339).
- Lüthi, S. and S. Wolter (2020). Are apprenticeships business cycle proof? *Swiss Journal of Economics and Statistics* 156(1), 1–11.
- Mohrenweiser, J. and U. Backes-Gellner (2010). Apprenticeship training: for investment or substitution? *International Journal of Manpower* 31(5), 545–562.
- Muehlemann, S., H. Dietrich, G. Pfann, and H. Pfeifer (2022). Supply shocks in the market for apprenticeship training. *Economics of Education Review* 86(C).
- OECD (2021). Pensions at a glance 2021: OECD and G20 indicators. OECD Publishing, Paris.
- OECD (2023). Main economic indicators. OECD Publishing, Paris.
- Oswald-Egg, M. E. and M. Siegenthaler (2021). Train drain? access to skilled foreign workers and firms' provision of training. *Labour Economics* 85, 102436.

- Ottaviano, G. and G. Peri (2012). Rethinking the effect of immigration on wages. *Journal of the European Economic Association* 10(1), 152–197.
- Riphahn, R. T. and M. Zibrowius (2016). Apprenticeship, vocational training, and early labor market outcomes – evidence from East and West Germany. *Education Economics* 24(1), 33–57.
- Roodman, D., M. Nielsen, J. G. MacKinnon, and M. D. Webb (2019). Fast and wild: Bootstrap inference in stata using boottest. *The Stata Journal* 19(1), 4–60.
- Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers* 3(2), 135–146.
- Vom Berge, P., C. Frodermann, A. Schmucker, and S. Seth (2021). Sample of integrated labour market biographies (siab) 1975-2019. FDZ-Datenreport 01/2021, Institute for Employment Research (IAB).
- Vuillemin, C. (2023). 'seit ich in der schweiz arbeite, schwelge ich im luxus.'. *SpiegelOnline*. 08.08.2023.
- Wenzelmann, F. and G. Schönfeld (2022). Kosten und Nutzen der dualen Ausbildung aus Sicht der Betriebe. Ergebnisse der sechsten BIBB-Kosten-Nutzen-Erhebung. Berichte zur beruflichen bildung, Federal Institute for Vocational Education and Training (BIBB).
- Wittek, B. H. and S. Muehlemann (2021). The consequences of eu eastern enlargement on human capital accumulation and wages in germany. Economics of Education Working Paper Series 184, University of Zurich, Department of Business Administration (IBW).

Appendix

Table A.1: Worker wages in Switzerland and Germany by ISCO-08 occupation in 1999-2002 in Euro in prices of 2000

Occupational Field	Switzerland	Germany
Managers	6,704.56 (5,625.24) [2,836]	3,044.47 (2,092.71) [19,401]
Professionals	5,718.79 (5,267.57) [6,394]	2,929.95 (1,865.70) [38,709]
Technicians and Associate Professionals	4,643.1 (4,220.43) [6,865]	2,738 (1,686.83) [42,886]
Clerical Support Workers	3,575.64 (3,446.68) [4,574]	2,811.97 (1,734.44) [12,640]
Services and Sales Workers	3,199.07 (2,949.34) [3,748]	2,124.83 (1,297.22) [9,460]
Skilled Agricultural, Forestry and Fishery Workers	2,493.55 (2,519.66) [465]	2,025.37 (1,018.28) [606]
Craft and Related Trades Workers	3,215.24 (3,267.02) [5,592]	2,753.92 (1,576.11) [20,377]
Plant and Machine Operators and Assemblers	3,554.22 (3,433.04) [1,754]	2,651.51 (1,606.26) [15,629]
Elementary Occupations	2,782.22 (3,044.49) [526]	2,405.58 (1,359.67) [28,084]

Note: Mean average wages of workers excluding apprentices by ISCO-08 first digit occupational field, Swiss wages converted from Swiss Francs to Euro using the 2002 average exchange rate, standard deviation in parentheses and number of observations in squared brackets. Source: SLFS and SIAB 1999-2002.

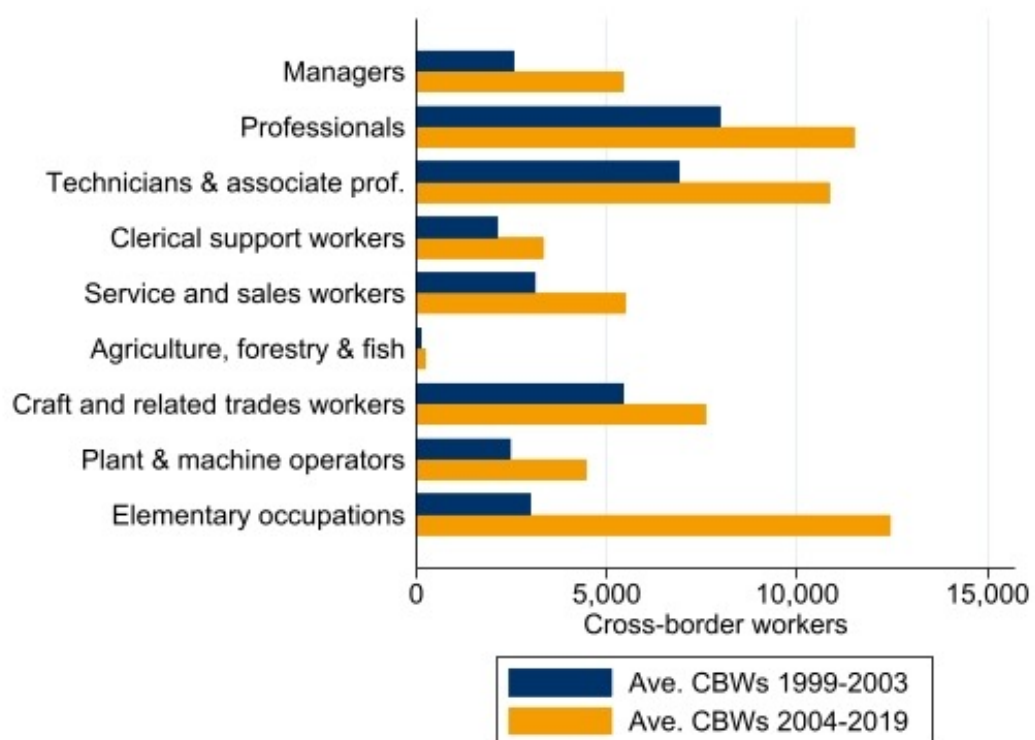


Figure A.1: Yearly average of registered cross border workers by occupation in the Transition Phase and the Free Movement Phase. Source: CCS 1999-2019

Table A.2: Comparison of district location by district location in 1998

	Border Region	Matched districts	West Germany
Mean population size	198,347 (54225.8)	257,254 (153,639.9)	409,948.2 (388,039.5)
Mean population density	357.9 (396.0)	433.7 (335.8)	1,027.4 (933.9)
Mean unemployment rate	6.970 (1.420)	10.468 (2.293)	10.422 (2.581)
Mean commuter balance	1.090 (17.242)	2.992 (24.644)	2.140 (29.970)
Mean number of firms	2,612.3 (693.4)	2,899.3 (1,639.9)	4,994.6 (5,128.5)
Share training firms	0.240 (0.019)	0.246 (0.030)	0.228 (0.042)
Mean employment level (FTE)	25,350.4 (8,011.1)	32,565.6 (17,592.1)	63,440.8 (72,442.4)
Mean number of apprentices	1,665.1 (412.8)	2,064.8 (1,034.9)	3,338.5 (3,282.6)
Share of apprentices	0.070 (0.009)	0.067 (0.013)	0.063 (0.017)
Share untrained workers	0.169 (0.028)	0.156 (0.030)	0.139 (0.026)
Share trained workers	0.764 (0.032)	0.766 (0.051)	0.783 (0.052)
Share workers with tertiary degree	0.100 (0.038)	0.110 (0.056)	.103 (0.049)
Share manufacturing workers	0.434 (0.115)	0.359 (0.114)	0.322 (0.139)
Mean log. Wages	83.7 (8.2)	82.4 (9.0)	85.9 (11.5)
Share censored wages	0.073 (0.031)	0.067 (0.037)	0.087 (0.052)

Note: Here, West Germany only considers municipalities that are located at least three hours by car away from the next Swiss city. Weighted by employment level of district. Source: BHP 1998.

Table A.3: Logistic regression results for the effect of the AFMP on the probability of firms to train

	Coefficient
$Transition \times \mathbb{1}(BR, d < 30)$	0.118* (0.041)
$Transition \times \mathbb{1}(BR, d \geq 30)$	-0.024 (0.041)
$Free \times \mathbb{1}(BR, d < 30)$	0.268*** (0.074)
$Free \times \mathbb{1}(BR, d \geq 30)$	0.205** (0.063)
Observations	188,848
Log likelihood	-77,945.093
LR chi2	2,516.27

Notes: Estimation following specification of equation (1), dependent variable: training in firm (Dummy-variable equal to 1 if firm trains apprentices), controls for firm fixed effects, year effects, and a NUTSII-regional time trend, excluding the city of Freiburg, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Source: BeH 1990-2019.

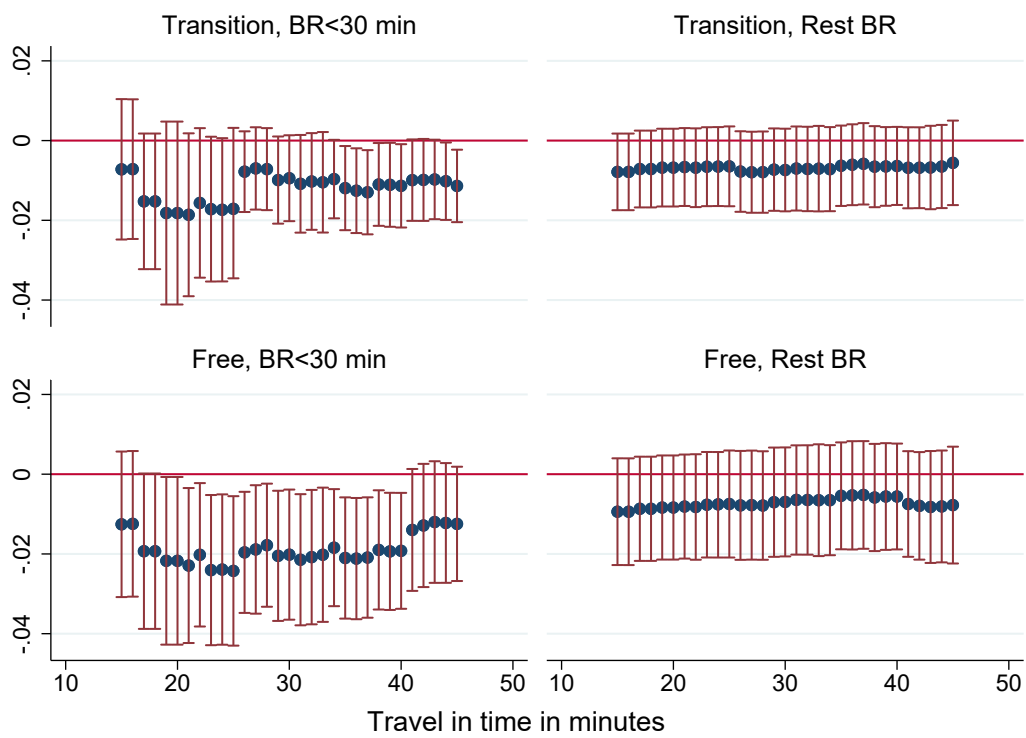


Figure A.2: Coefficients of the effect on the apprentice wage when estimating Equation 5 using different cut-off travel times, excluding the city of Freiburg, 95%-confidence intervals based on robust standard errors adjusted for 61 clusters at the level of municipal associations. Source: BeH 1990-2019

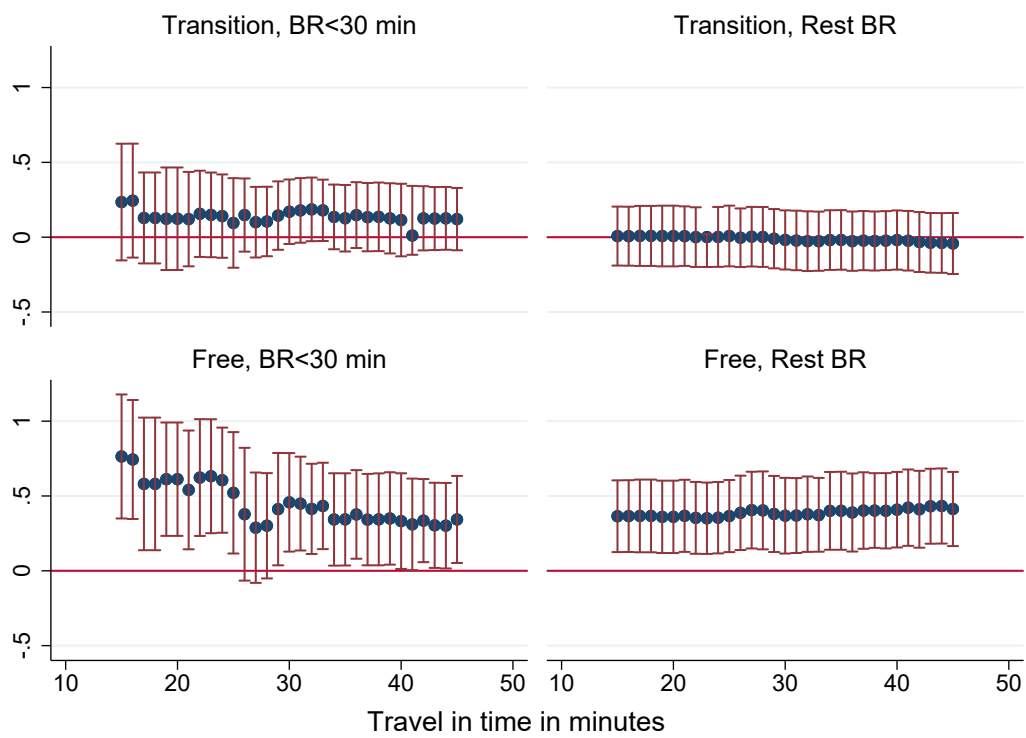


Figure A.3: Coefficients of the effect on the apprentice numbers in firms when estimating Equation 5 using different cut-off travel times, excluding the city of Freiburg, 95%-confidence intervals based on robust standard errors adjusted for 61 clusters at the level of municipal associations. Source: BeH 1990-2019

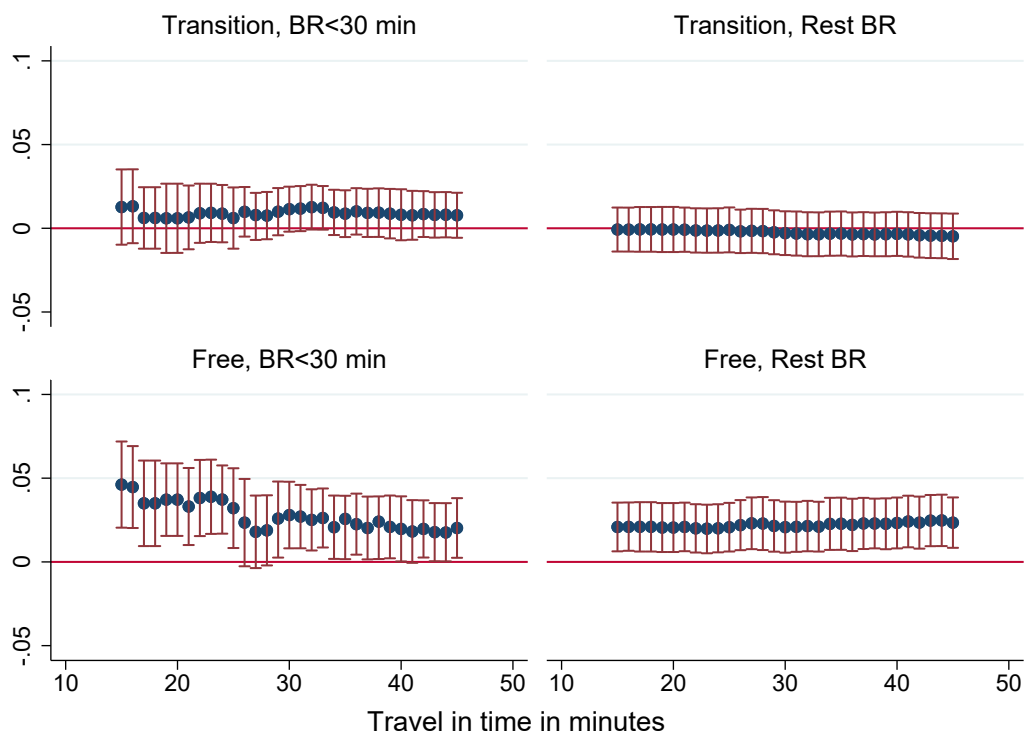


Figure A.4: Coefficients of the effect on the training probability when estimating Equation 5 using different cut-off travel times, excluding the city of Freiburg, 95%-confidence intervals based on robust standard errors adjusted for 61 clusters at the level of municipal associations. Source: BeH 1990-2019

Table A.4: Robustness Check: Results using travel time to the next Swiss city as a continuous treatment measure.

	(1) Apprentice wage	(2) Number of apprentices	(3) Probability of training
<i>Transition</i> × Travel time	0.005* (0.002) [0.026]	-0.022 (0.042) [0.595]	-0.001 (0.003) [0.615]
<i>Free</i> × Travel time	0.006* (0.003) [0.026]	-0.124* (0.053) [0.044]	-0.008* (0.003) [0.038]
Observations	130,335	332,138	332,138
R2	0.760	0.684	0.560

Notes: Estimation following specification of equation (1), dependent variables: (1) logarithm of mean apprentice wages, (2) logarithm of the mean-averaged number of apprentices, and (3) training in firm (Dummy-variable equal to 1 if firm trains apprentices), controls for firm fixed effects, year effects, and a NUTSII-regional time trend, excluding the city of Freiburg. Robust standard errors adjusted for 50 clusters at the level of municipal association in parentheses and wild boot strap p-values in brackets, *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Source: BeH 1990-2019.

Table A.5: Robustness Check: Results including the city of Freiburg.

	(1) Apprentice wage	(2) Number of apprentices	(3) Probability of training
$Transition \times \mathbb{1}(BR, d < 30)$	-0.007 (0.006) [0.198]	0.164 (0.106) [0.170]	0.011 ⁺ (0.007) [0.122]
$Transition \times \mathbb{1}(BR, d \geq 30)$	-0.010 ⁺ (0.005) [0.106]	-0.022 (0.095) [0.830]	-0.003 (0.006) [0.663]
$Free \times \mathbb{1}(BR, d < 30)$	-0.016 ⁺ (0.009) [0.108]	0.441** (0.158) [0.008]	0.028** (0.009) [0.008]
$Free \times \mathbb{1}(BR, d \geq 30)$	-0.011 (0.008) [0.176]	0.338** (0.120) [0.010]	0.020** (0.007) [0.014]
Observations	136,448	350,136	350,136
R2	0.759	0.685	0.560

Notes: Estimation following specification of equation (1), dependent variables: (1) logarithm of mean apprentice wages, (2) logarithm of the mean-averaged number of apprentices, and (3) training in firm (Dummy-variable equal to 1 if firm trains apprentices), controls for firm fixed effects, year effects, and a NUTSII-regional time trend. Robust standard errors in parentheses adjusted for 62 municipal association clusters, wild bootstrap p-values in brackets, *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Source: BeH 1990-2019.

Table A.6: Robustness Check: Results excluding the city of Freiburg and Waldshut-Tiengen

	(1) Apprentice wage	(2) Number of apprentices	(3) Probability of training
$Transition \times \mathbb{1}(BR, d < 30)$	-0.009 ⁺ (0.005) [0.096]	0.159 (0.109) [0.158]	0.011 (0.007) [0.124]
$Transition \times \mathbb{1}(BR, d \geq 30)$	-0.008 (0.005) [0.176]	-0.015 (0.10) [0.902]	-0.003 (0.007) [0.713]
$Free \times \mathbb{1}(BR, d < 30)$	-0.020* (0.008) [0.032]	0.430* (0.168) [0.016]	0.026* (0.010) [0.014]
$Free \times \mathbb{1}(BR, d \geq 30)$	-0.007 (0.007) [0.399]	0.364** (0.130) [0.014]	0.020* (0.008) [0.028]
Observations	128,901	329,453	329,453
R2	0.760	0.683	0.559

Notes: Estimation following specification of equation (1), dependent variables: (1) logarithm of mean apprentice wages, (2) logarithm of the mean-averaged number of apprentices, and (3) training in firm (Dummy-variable equal to 1 if firm trains apprentices), controls for firm fixed effects, year effects, and a NUTSII-regional time trend, excluding the city of Freiburg and Waldshut-Tiengen. Robust standard errors in parentheses adjusted for 60 municipal association clusters, wild bootstrap p-values in brackets, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Source: BeH 1990-2019.

Table A.7: Effect of the AFMP on the apprentice retention rate by tenure after apprenticeship

	Retention in		
	year 1	year 2	year 3
<i>Transition</i> × $\mathbb{1}(BR, d < 30)$	-0.015 (0.011) [0.224]	-0.007 (0.011) [0.583]	-0.006 (0.010) [0.595]
<i>Transition</i> × $\mathbb{1}(BR, d \geq 30)$	-0.015 (0.011) [0.184]	-0.014 (0.011) [0.196]	-0.009 (0.010) [0.439]
<i>Free</i> × $\mathbb{1}(BR, d < 30)$	-0.035* (0.015) [0.050]	-0.031** (0.014) [0.020]	-0.028* (0.011) [0.014]
<i>Free</i> × $\mathbb{1}(BR, d \geq 30)$	-0.022 (0.015) [0.162]	-0.032* (0.014) [0.024]	-0.025* (0.011) [0.040]
Observations	130,335	130,335	130,335
R2	0.357	0.349	0.344

Notes: Estimation following specification of equation (1), dependent variable: Retention rate calculated as the number of apprentices still working at the firm after 1 year (Column 1), 2 years (Column 2), and 3 years (Column 3), considering only apprentices that have completed an at least 2-year apprenticeship period within the firm, rates are a mean-average over three years, controls for firm fixed effects, year effects, and a NUTSII-regional time trend, excluding the city of Freiburg. Robust standard errors in parentheses adjusted for 61 municipal association clusters, wild bootstrap p-values in brackets, *** p<0.001, ** p<0.01, * p<0.05, + p<0.1. Source: BeH 1990-2019.

Table A.8: Occupational composition of apprenticeships in the BR over regulation phases.

Occupational Field	Pre-policy 1990-1998	Transition 1999-2003	Free 2004-2019	Mean wage 1996-1998
Agriculture, forestry, farming, and gardening	0.018 (0.130)	0.020 (0.140)	0.018 (0.133)	21.85 (9.79)
Production of raw materials and goods, and manufacturing	0.344 (0.475)	0.339 (0.473)	0.363 (0.481)	25.77 (7.84)
Construction, architecture, surveying and technical building services	0.103 (0.304)	0.093 (0.290)	0.087 (0.266)	28.41 (11.11)
Natural sciences, geography and informatics	0.013 (0.114)	0.014 (0.118)	0.016 (0.125)	28.55 (7.84)
Traffic, logistics, safety and security	0.010 (0.101)	0.014 (0.117)	0.018 (0.134)	31.28 (16.40)
Commercial services, trading, sales, the hotel business and tourism	0.141 (0.348)	0.156 (0.363)	0.133 (0.340)	25.04 (7.24)
Business organisation, accounting, law and administration	0.141 (0.348)	0.144 (0.351)	0.157 (0.364)	28.69 (10.41)
Health care, the social sector, teaching and education	0.215 (0.411)	0.206 (0.405)	0.200 (0.400)	30.61 (12.74)
Philology, literature, humanities, social sciences, economics, media, art, culture, and design	0.014 (0.116)	0.014 (0.118)	0.007 (0.088)	31.67 (18.45)
Observations	81,585	39,705	100,631	26,215

Note: Columns 1-3 show the average share of apprenticeships in BR by first digit occupational group of the Classification of Occupations 2010, Column 4 shows the average apprentice wage in these occupations, excluding the city of Freiburg, standard deviation in parentheses. Source: BeH 1990-2019.

Table A.9: Effect of the AFMP on skilled workers in firms.

	Employment	Wage All	Wage Incumbents
$Transition \times \mathbb{1}(BR, d < 30)$	-0.044* (0.017) [0.036]	-0.025** (0.009) [0.018]	0.000 (0.009) [0.976]
$Transition \times \mathbb{1}(BR, d \geq 30)$	-0.003 (0.010) [0.822]	-0.017* (0.007) [0.064]	-0.005 (0.006) [0.467]
$Free \times \mathbb{1}(BR, d < 30)$	-0.033 (0.021) [0.152]	0.000 (0.013) [0.972]	0.027 (0.017) [0.214]
$Free \times \mathbb{1}(BR, d \geq 30)$	0.012 (0.017) [0.563]	-0.005 (0.011) [0.749]	0.000 (0.009) [0.980]
Observations	313,837	284,512	219,343
R2	0.885	0.800	0.854

Notes: Estimation following specification of equation (1), dependent variables: (1) logarithm of number of employed skilled workers, (2) logarithm of the full-time mean wage of all skilled workers and (3) skilled workers that have been already employed at the firm at least one year before the reform announcement, controls for firm fixed effects, year effects, and a NUTSII-regional time trend, excluding the city of Freiburg. Robust standard errors in parentheses adjusted for 61 municipal association clusters and wild bootstrap p-values in brackets, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, + $p < 0.1$. Source: BeH 1990-2019.

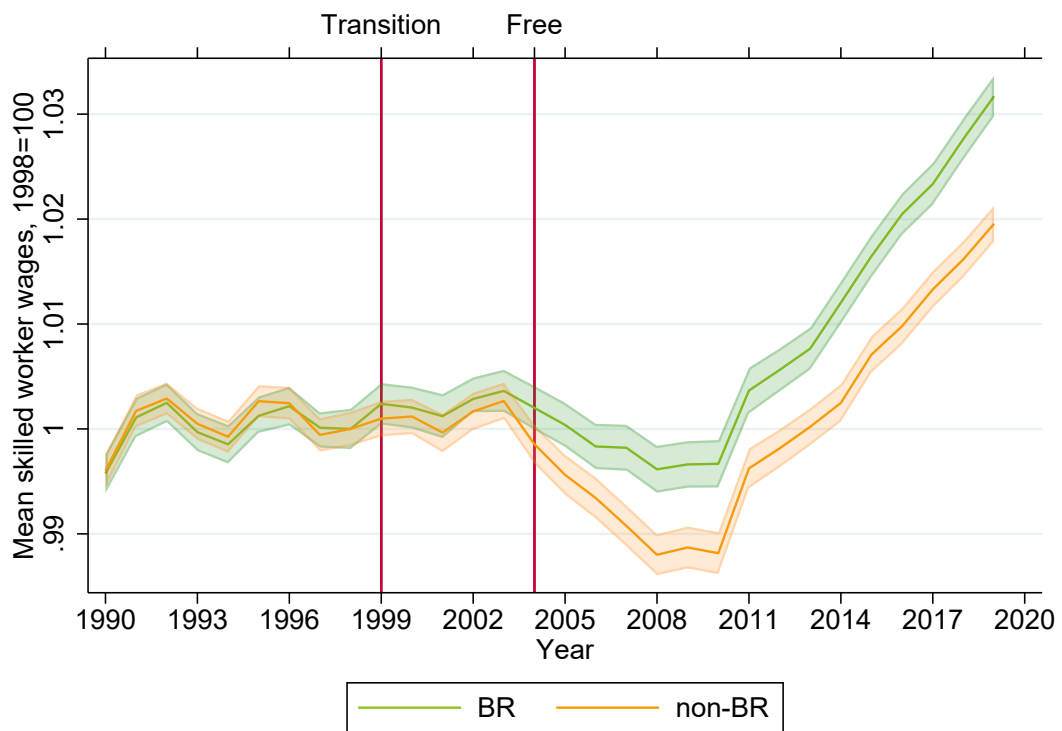


Figure A.5: Development of mean skilled worker wages in the border region and the control regions with 95% confidence intervals based on robust standard errors adjusted for 61 clusters at the level of municipal associations. Source: SIAB 1990-2019.