SFB 823

The effects of reforming a federal employment agency on labor demand

Kornelius Kraft, Alexander Lammers

Nr. 28/2019

Discussion Paper

SFB 823

The Effects of Reforming a Federal Employment Agency on Labor Demand

Kornelius Kraft[†]   Alexander Lammers[‡]

May 2021

Abstract

In this paper we report the results of an empirical study on the employment growth effects of a policy intervention, explicitly aimed at increasing placement efficiency of the Federal Employment Agency in Germany. We use the Hartz III reform in the year 2004 as an exogenous intervention that improves the matching process and compare establishments that use the services of the Federal Employment Agency with establishments that do not use the placement services. Using detailed German establishment level data, our difference-in-differences estimates reveal an increase in employment growth among those firms that use the agency for their recruitment activities compared to non-user firms. After the Hartz III reform was in place, establishments using the agency grew roughly two percentage points faster in terms of employment relative to non-users and those establishments achieve an increase in the proportion of hires. We provide several robustness tests using for example inverse-probability weighting to additionally account for differences in observable characteristics. Our paper highlights the importance of the placement service on the labor demand side, in particular on the so far overlooked establishment level.

JEL Classification: J23, J64, J68

Keywords: Hartz III reform, Federal Employment Agency, matching efficiency, employment growth, difference-in-differences

Acknowledgments: We thank the German Research Foundation (DFG) for its financial support of the Collaborative Research Center "Statistical modeling of non-linear dynamic processes" (SFB 823, Project A5). We also thank the German Federal Employment Agency (BA) and the Institute for Employment Research (IAB) for providing the data. We received very helpful comments and suggestions from the participants of the international workshop on vacancies, hiring and matching at the IAB in Nuremberg. Any remaining mistakes are our own.

[†]TU Dortmund University, Faculty of Business and Economics, Vogelpothsweg 87, D-44227 Dortmund, ZEW Mannheim, IZA Bonn and KU Leuven, Belgium. Email: Kornelius.Kraft@tu-dortmund.de.

[‡]Corresponding author: TU Dortmund University, Faculty of Business and Economics, Vogelpothsweg 87, D-44227 Dortmund. Phone: (+49)231 755-3156. Email: Alexander.Lammers@tu-dortmund.de.

1 Introduction

The functioning of frictional labor markets with imperfect and asymmetric information largely depends on the efficiency of the matching process between unemployed and vacancies (e.g. Petrongolo & Pissarides 2001; Mortensen & Pissarides 1994). In recent years, economies, firms and employees have faced serious challenges within the labor market. The automation of tasks within jobs has accelerated and an increasing proportion of jobs is at risk of being replaced by advanced technologies such as algorithms or robots (e.g. Acemoglu & Restrepo 2018; Frey & Osborne 2017; Brynjolfsson & McAfee 2014). At the same time, labor markets in several European countries are further challenged by large inflows of workers that need to be efficiently integrated into the labor market (Battisti et al. 2019). The matching of workers to vacancies is therefore becoming an increasingly urgent task, while being extremely demanding and highly important in terms of government spending. While the effects on workers of an increase in the labor market matching efficiency such as unemployment duration or satisfaction have been studied broadly, it has not generally been analyzed in detail from the labor demand perspective. The literature, however, provides evidence that matching rates and the filling of vacancies are rather firm-specific (Kaas & Kircher 2015; Davis et al. 2013). Moreover, the placement process depends, among other factors, on the effectiveness of labor market institutions such as the Federal Employment Agency (FEA).

In this paper we investigate an important policy reform that was explicitly framed at improving the employment agency in terms of job matching efficiency in Germany. During the first years of the 21st century, various labor market reforms were implemented in Germany. This was embedded in the so-called Hartz reform package, successively implemented between the years 2003–2005, and consists of the reforms Hartz I–IV. Our focus is on the Hartz III reform, which became effective on January 1, 2004. We exploit this exogenous policy intervention aiming to improve the efficiency of the FEA and investigate whether establishments using the FEA for their job recruitment benefit from an improvement in the internal restructuring of the FEA. We measure this improvement in terms of employment creation of establishments using the placement services, compared to establishments not using the placement services.

We use highly detailed establishment level information for the years 2000 to 2008 from the German IAB Establishment Panel provided by the Institute for Employment Research (IAB). From this data set we create a sample of 14,658 establishment-year observations. We apply difference-in-differences estimation in which the establishments using the FEA constitute the treatment group and establishments which do not use the placement services the control group. Using this estimation framework allows us to: (1) estimate the causal link between reforming a federal agency and employment creation of establishments and (2) account for macroeconomic common shocks, for example that the Hartz reforms were implemented during an expansionary time in Germany (e.g. Bradley & Kügler 2019).[1] Robustness tests are provided in Section 5 in which we explicitly look at selection effects for choosing the employment agency as a recruitment channel in the first place using inverse-probability-weighting (IPW) with different specifications.

In terms of employment creation, we look at the share of new hires as the ratio of hirings to total employment as well as employment growth. The unweighted regression results indeed provide evidence of positive reform effects for establishments using the placement services relative to establishments which do not use the FEA. The effects are in the magnitude of a roughly 2-percentage-point increase in the share of hires. According to our estimation results, the reform of the federal employment agency led to higher growth of employment by roughly 3 percentage points. The weighted regression results are slightly smaller.

Our paper contributes to the microeconomic literature on matching efficiency as well as to the literature on the evaluation of the Hartz reforms.[2] There are also macroeconomic studies that examine the impact of the Hartz III reform, for example by considering unemployment duration or aggregate flows into and out of unemployment. In contrast,

[1] Regarding macroeconomic shocks, for example Davis & Haltiwanger (1992) report that significant job creation and destruction coexist in all phases of the business cycle.

[2] A comprehensive summary of micro-evaluation studies regarding the Hartz reforms can be found for example in Akyol et al. (2013).

we use a microeconomic approach and examine the labor demand side and the effects of the reform on the establishment level. We therefore examine whether the behavior of establishments has actually changed since the reform is in place, which we measure in terms of employment growth and hiring rates.

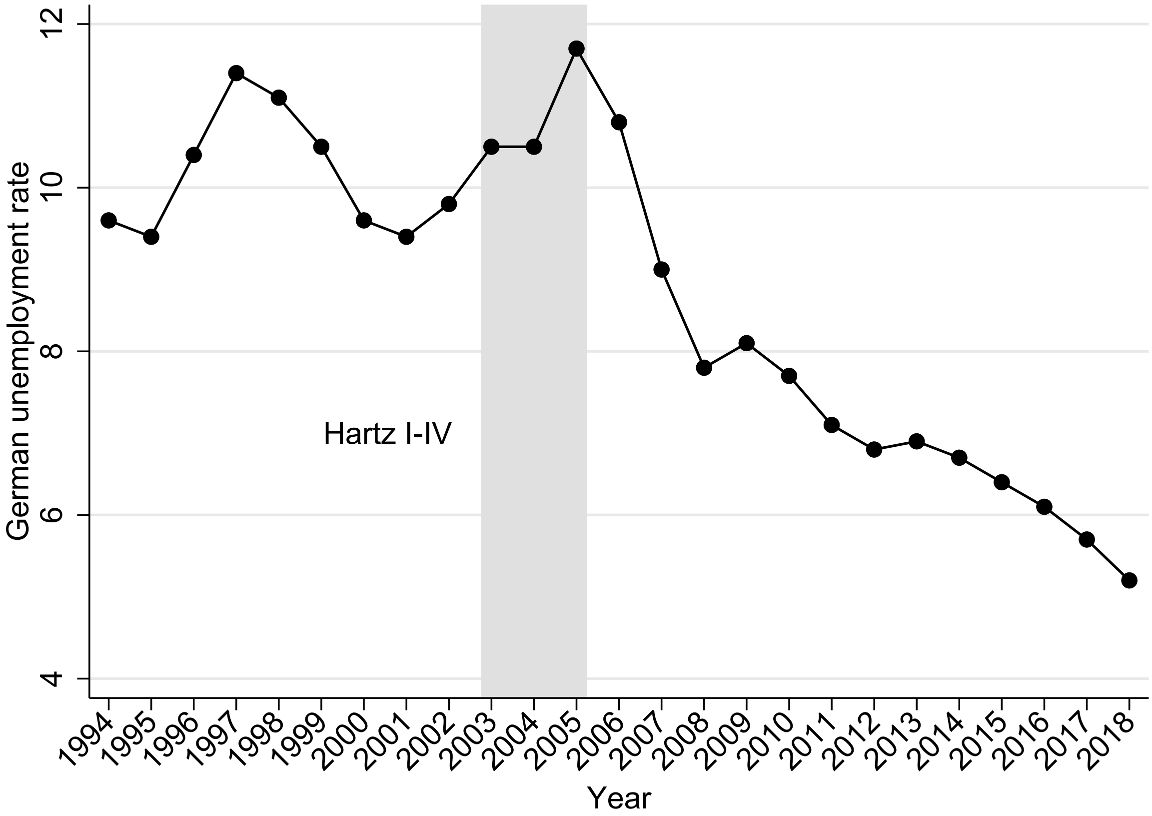
The paper proceeds as follows. In Section 2 we provide a literature review on micro- and macroeconomic studies regarding matching efficiency, particularly in the context of the Hartz legislation. Section 3 then provides theoretical arguments for the connection between use of the Federal Employment Agency, matching efficiency and employment growth. Our empirical investigation is provided in Section 4 and robustness tests are provided in Section 5. Finally, Section 6 draws a conclusion and provides policy implications.

2 Facts of the Hartz reform legislation

2.1 Overview of German unemployment

The early 2000s in Germany were characterized by a high and persistent unemployment rate at roughly 10 percent, peaking at 11.1 percent in 2005 (Dustmann et al. 2014). Figure 1 provides an overview of the German unemployment rate in which a substantial drop in the unemployment rate is discernible after the Hartz reforms were implemented. Figure A.1 in the Appendix provides a more nuanced view by dividing the analysis into western and eastern Germany. Besides high and persistent unemployment, the motivation for the introduction of labor market reforms has been further strengthened by the so-called placement scandal triggered by the Federal Employment Agency in the year 2002 (Fleckenstein 2008). In this scandal the FEA manipulated statistics, significantly exaggerating the numbers of successfully placed job seekers. These were the two main reasons leading to the appointment of the Hartz Commission on February 22nd, 2002 to suggest labor market reforms. The Hartz Commission, named after the chairman of the commission Peter Hartz, consisted out of 15 experts from industry, politics and academia. The commission published the suggestions for labor market reforms in August 2002 which finally led to the Hartz reform package.

Figure 1: German unemployment rate over time



Notes: The figure shows the trend in the German unemployment rate over time. Data is provided by the Federal Employment Agency time series: "Unemployment over time". The unemployment rate has its peak at 11.1 percent in the year 2005 after which the rate remarkably falls. The gray shaded area marks the step-by-step introduction of the Hartz reform package from 2003 to 2005. Hartz I and II became effective on January 1, 2003, Hartz III on January 1, 2004 and finally Hartz IV on January 1, 2005. For the unemployment rate divided between eastern and western Germany, see Figure A.1 in the Appendix.

The Hartz reforms are divided into four packages, which were introduced successively and affected almost all aspects of the German labor market. Since the reforms came with an evaluation mandate from the government, several empirical studies have been conducted regarding the Hartz Reforms; Jacobi & Kluve (2007) for example provide an overview of this. Hartz I and II were introduced and became effective on January 1, 2003 and aimed at improving labor market flexibility through "Mini-Jobs" legislation. In particular, Hartz I facilitated easier hiring of temporary workers by lifting employment restrictions and, in addition to that, further training for employees was subsidized by vouchers. Hartz II reorganized marginal employment by raising the tax-free earnings threshold from 325 to 400 EUR tax-free income per month.

Hartz III became effective on January 1, 2004 and had the primary objective of increasing the internal efficiency of the Federal Employment Agency. The most important change was the reorientation of the agency towards a customer-oriented service facility, in which all unemployment claims were handled by one designated case worker. The Hartz III reform transformed the employment agency from a centralized budgeting system to a more management-by-objectives system with clearly defined tasks and goals (e.g. Akyol et al. 2013). Moreover, the contact time between case workers and job seekers was increased and different advisory services were introduced for the short- and long-term unemployed. Furthermore, so called „Job Centers“ Job Centers were implemented with the further aim of improving the placement process by enhancing competition among them. The main goal was to reduce frictions and improve the matching efficiency between employers and job seekers.

Finally, Hartz IV came into effect on January 1, 2005 and was aimed at shortening the duration of the higher unemployment benefit ALG I (*Arbeitslosengeld*) paid to the newly unemployed. The reform therefore reduces the long-term unemployment benefit ALG II. Furthermore, sanctions are now implemented to increase the incentives of more active labor market support. To date, the Hartz IV reform is one of the most extensive and controversially discussed labor market reforms in Germany.

2.2 Related literature

Regarding matching efficiency, there is a burgeoning supply of literature evaluating the Hartz reforms in the past decade. With respect to Hartz II, for example, Bradley & Kügler (2019) found an increase in mini-job workers from 13 percent in 2003 to 16 percent in 2006. Dlugosz et al. (2014) investigate the Hartz IV reform and show that the reduction of unemployment benefit entitlement provides incentives for older workers to remain employed. In a similar vein, Krebs & Scheffel (2013) use a calibrated model to simulate the effects of Hartz IV which reduced structural unemployment by 1.4 percentage points. In combination with Hartz I–III, the aggregated effect is a 1.5 percentage point reduction in structural unemployment. Gehrke et al. (2019) find positive labor market performance shocks caused by the Hartz reforms. They argue that these reforms are the main driver for good performance during the great financial crisis in Germany between 2008 and 2009. With respect to the relevance of placement services, almost 50 % of all vacancies in Germany are registered at the Federal Employment Services. Moreover, the literature often finds that, compared to the private market, applicants sent by the employment agency are usually less suited for the job and thus firms pay lower wages for these applicants (e.g. Holzner & Watanabe 2015).[3] Pellizzari (2010) exploits a policy intervention in the Italian employment and recruitment services, which aimed at making the recruitment services more competitive. He finds higher wages for employees being matched via more efficient employment agencies. Using a synthetic control method, Ehrich et al. (2018) find that the Hartz reforms raised labor force participation, specifically among women and older workers.

While the studies discussed so far examine other effects of the Hartz reforms, more closely related to our research are some recent macroeconomic studies which consider matching efficiency. For example, Stops (2016) estimates parameters of macroeconomic matching functions before, during and after the Hartz reforms. He finds that matching productivity increased during all reform stages even after controlling for the business cycle. Fahr & Sunde (2009) show that the Hartz reforms accelerate outflows from unemployment to employment after the Hartz III reform had been implemented, in which the effects are more pronounced for eastern Germany. Launov & Wälde (2016) structurally estimate the reform effect of an increase in matching effectiveness on the unemployment rate. They provide evidence that the reorganization of the FEA is responsible for a .69-.88 percentage point decline of the equilibrium unemployment rate.[4] Moreover, Launov & Wälde (2016) highlight an unemployment paradox. A more effective FEA placement service for long-term unemployed workers might crowd out private search effort since unemployed anticipate an increase in placement probability. Klinger & Rothe (2012) also find increased matching efficiency by roughly 10 percent using simultaneous stock-flow

[3] Holzner & Watanabe (2015) also point out that more efficient Federal Employment Services might crowd out private search effort. This result is also found, for example, by Launov & Wälde (2016).

[4] This is a decline of about 17.7 to 22.5 % in post-reform unemployment. Hartz IV, however, is only responsible for a 4.6 to 5.1 % post reform unemployment decline (Launov & Wälde 2016).

matching functions for short-term and long-term unemployed. The result is supported by Klinger & Weber (2016), who find an extraordinary increase in matching efficiency after 2005. Hartung et al. (2018) argue that instead of an increased hiring rate, lower separation rates explain the decline in unemployment after the Hartz reforms. Recently, Bauer & King (2018) use a reallocation model to investigate the effects of the reforms. They found that the reforms significantly reduce reallocation costs and therefore unemployment in Germany in the long run.

The literature also provides evidence from other countries. For example, with regard to an increase in the duration of unemployment benefit, Le Barbanchon (2016) does not find any effects on the matching quality for France. Liechti (2020) shows for Switzerland, that recommendation from an employment agency can act as a substitute for social contacts. In a similar context, Horton (2017) considers the effect of algorithmic recommendations for employers. He finds that such recommendations are very effective for hiring, especially when firms are faced with a small pool of applicants. These results have important policy implications since it may be a good strategy to improve social connections between job seekers and employers.

Summing up, most of the reviewed literature focuses either (i) on aggregated effects, the reduction of unemployment duration and equilibrium effects or (ii) on effects on the level of unemployed workers. Explicit microeconomic studies with a focus on the Hartz III legislation and its impact on the labor demand side in terms of employment creation, in particular on the establishment level, are missing from this literature.

3 Job matching and employment growth

Labor market institutions such as the Federal Employment Agency exert a strong influence on the job matching process. Above all, the agency is responsible for bringing together supply and demand, i.e. the unemployed in search of a job and employers who post vacancies. A match is characterized by the placement of an unemployed person in a vacancy, in which the efficiency is determined by the matching function (e.g. Davis et al. 2013; Petrongolo & Pissarides 2001).[5] Within this process the Federal Employment Agency provides job search assistance which helps unemployed workers to find suitable jobs and monitors the search effort of unemployed people. Following the Hartz III reform, the efficiency of job placement has been greatly increased. The employment agency shifted from a centralized budgeting system to a more management-by-objectives system in which clearly defined tasks and goals are defined (Akyol et al. 2013). The contact time per unemployed person was increased and 'Job Centers' were established with the aim of improving the placement process further. With regard to these considerations, the Hartz III reforms can be considered a positive technological shock for the matching production function of the Federal Employment Agency (e.g. Petrongolo & Pissarides 2001).

Following the restructuring of the agency, unemployed workers are more closely monitored, in many cases better motivated and thus more suitable for the job market. Moreover, the agency may help employers to better overcome information asymmetries by placing workers in occupations that fit their qualifications. In a similar vein, Marinescu & Rathelot (2018) show that geographic mismatch is a potential driver for unemployment. Bauer & King (2018) argue that a more efficient employment agency can improve the placement results, because with their assistance the employees are also made aware of jobs outside of their former profession. The result is a reduction in mismatch caused by imperfect labor mobility.

A more efficient employment agency therefore reduces search and recruitment costs for employees but also for establishments. The reduction in search costs is associated with an increase in productivity since workers and establishments can consider potential matches more efficiently (Autor 2001; Pissarides 1990). Having access to more capable job candidates due to an efficient search channel leads to better matches, which may improve labor productivity and reduce the need for further training activities. Bryson & Nurmi (2011) point out in this context that specific job-related tasks can be performed more efficiently, resulting in a competitive advantage and in employment growth. Ultimately,

[5] According to various job search models, employers post vacancies to attract potential job seekers. The matching function then links the combination of job seekers and job vacancies and produces new hires (e.g. Davis et al. 2013).

better matches between employers and employees lower search and recruitment costs for employers, thus facilitating the process of job creation (e.g. Blasco & Pertold-Gebicka 2013; Pissarides 1990). This line of reasoning, that lowering search costs is associated with higher productivity is well established in the labor market search theory (e.g. Autor 2001; Pissarides 1990). Moreover, Blasco & Pertold-Gebicka (2013) note that firms' performance in the short run might be reduced due to adaptation costs. However, long-run effects might indeed be positive.

To sum up, a reduction in search costs due to the Federal Employment Agency's more efficient placement process leads to better matches, reduces the necessity for further training activities and ultimately increases productivity (e.g. Autor 2001; Pissarides 1990). Ultimately, this mechanism increases the competitiveness of the benefiting establishment and facilitates job creation. Whether this hypothesis really applies is the subject of the following empirical test, in which we test whether the Hartz III reform is indeed associated with employment growth among those establishments that actually use the placement services.

4 Empirical Evidence

4.1 IAB Establishment Panel

To examine the effect of the Hartz III reform on the establishment level, we use data from the German IAB Establishment Panel provided by the Institute for Employment Research (IAB).[6] This panel has been conducted on an annual basis since 1993 in western Germany and since 1996 in eastern Germany and surveys roughly 16,000 establishments per year. The panel is designed to lead to a representative sample for Germany which is explicitly analyzed, for example by Bossler et al. (2018). The questionnaire asks about a wide variety of establishment characteristics including the use of the Federal Employment Agency as a recruitment channel for establishments. This information is crucial for our

[6] For more details regarding the sampling methodology see for example Bossler et al. (2018); Ellguth et al. (2014); Fischer et al. (2009).

identification strategy and is available for the survey years 2000 to 2008 in which we are able to create a sample of 14,658 establishment-year observations. Descriptive statistics are provided in Table 2.

4.2 Treatment and control group

To divide establishments into a treatment (employment agency user) and control group (non-user) we use information on whether establishments use the Federal Employment Agency as a recruitment channel. More precisely, we have information on vacancies[7] reported to the agency. We utilize this information to construct a treatment indicator which takes unit value for firms which continuously report vacancies greater than zero to the employment agency and additionally report vacancies at all for every sample year and zero otherwise. Establishments might anticipate a more effective placement process and therefore start to use the placement services. Our treatment indicator, however, is exogenously constructed before the reform was in place and thus we consider establishments that do not change their job search behavior, i.e. do not switch between FEA and private agents. We therefore assign establishments to the treatment and control group before the treatment occurs in 2004. On the other side, the control group consists of establishments which report zero vacancies to the Federal Employment Agency and also report vacancies greater than zero which ensures that both groups are comparable. In doing so, we are able to distinguish establishments between the year 2000 and 2008 that are directly affected by an improvement of the placement service and establishments that are not.[8] An overview regarding the distribution of establishments using the federal placement services compared to those which are not, is provided in Table 1.

It becomes evident, that smaller establishments tend to be more prone to other recruitment channels and establishments employing more workers tend to also rely more on the federal employment agency as a recruitment channel. Moreover, Table 1 shows the

[7] The question from the questionnaire reads, above all: "How many vacancies have you planned to be filled immediately? [. . .] How many of these vacancies are registered with the employment office?"

[8] See for example Blasco & Pertold-Gebicka (2013) and Hud & Hussinger (2015) for a similar approach in classifying the treatment and control group.

Table 1: Raw distribution of vacancies and establishment size

	Treatment group				Control group			
	Observations		#Vacancies		Observations		#Vacancies	
	N	%	all	FEA	N	%	all	FEA
Employees	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
1-19	1,717	20.57	1.74	1.68	1,928	30.56	1.47	0.00
20-49	1,343	16.09	3.52	3.39	1,267	20.08	2.07	0.00
50-199	2,239	26.82	6.34	5.98	1,644	26.05	3.30	0.00
200-499	1,618	19.38	8.46	7.50	849	13.45	5.40	0.00
500+	1,431	17.14	20.37	17.66	622	9.86	17.07	0.00
Total	8,348	100			6,310	100		

Notes: This table shows the distribution of establishments through size categories. The control group consists of establishments that do not use the Federal Employment Agency (FEA), which is identified as reported in Section 4.2. The treatment group consists of establishments that use the placement services and thus are affected by the Hartz III reform which became effective in January 1st 2004. Data from the IAB Establishment Panel, waves 2000–2008, with an overall sample size of $N = 14,658$ establishment-year observations.

reported vacancies from the treatment and control group. As shown, the treatment group reports slightly more vacancies in every establishment size category. Most important, however, for the group definition regarding treatment and control group are columns (4) and (8) in Table 1 which shows the vacancies among both groups which are reported to the Federal Employment Agency. Interestingly, the rates of vacancies in both groups are very similar, which is favorable for a comparison. By definition, reports of vacancies to the FEA are zero for the control observations. The treatment observations report quite a high ratio of their vacancies to the employment agencies.

4.3 Methodology

Dependent Variables. For the difference-in-differences specification, we create the following dependent variables. As is standard in the literature (e.g. Chodorow-Reich 2014; Davis et al. 2013), we compute a symmetric employment growth rate as the difference in the number of employees E_{it} in establishment i at year t and year $t-1$, divided by the average of employees in both years:

$$g_{it} = \frac{E_{it} - E_{it-1}}{(E_{it} + E_{it-1})/2} \tag{1}$$

The trend of employment growth as calculated in Equation (1) over time between the treatment and control group, is shown in Figure 3. The employment growth calculated this way is quite convenient since the rate is bounded in the range $[-2, 2]$ and furthermore can accommodate employee entries and exits, which is explicitly helpful to limit the influence of outliers in employment growth.[9] Second, we use the share of hires in relation to existing employment as proposed for example by Gralla & Kraft (2018) which is defined as the number of hires h_{it} in the year $t+1$ divided by the number of employees E_{it} in establishment i at year t. The trend of the share of hires, as calculated in Equation (2) over time between the treatment and control group, is shown in Figure 2.

$$sh_{it} = \frac{100 * h_{it+1}}{E_{it}} \tag{2}$$

We expect this share to be positively affected by an increased employment service performance in the treatment group relative to the control group since the Hartz III reform has been in place. We consider hirings and do not differentiate between employees who were previously unemployed or in employment (job-to-job transitions). For a similar approach see, for example, Blasco & Pertold-Gebicka (2013), who consider new hires stemming from the pool of unemployed, and the procedure applied by Bauer & King (2018), who consider job-to-job transitions. We model the joint movement of job-to-job seekers and job seekers who are currently unemployed.

Estimation Framework. To measure the effects of an increase in placement service efficiency we rely on a difference-in-differences estimation strategy to measure whether establishments tend to exhibit a higher employment growth. For the share of hires as the

[9] In this context, see for example Chodorow-Reich (2014); Brändle & Goerke (2018); Bryson (2004); Wooden & Hawke (2000) for a similar specification of employment growth, however in different economic contexts. Furthermore, this measure has the property of being approximately normally distributed.

dependent variable, we apply a corner solution model estimated by a heteroscedastic tobit model to take the fraction of non-hiring establishments into account.[10] For robustness and to allow for fixed effects (FE), we also apply OLS models. We estimate the following specification:

$$
\begin{aligned}
y_{it} = \alpha &+ \beta_1 FEAuser_i + \beta_2 HartzIII_t + \tau HartzIII_t \times FEAuser_i \\
&+ \beta_m X_m + \gamma_t + \rho_i + \lambda_i + \varepsilon_{it}
\end{aligned} \tag{3}
$$

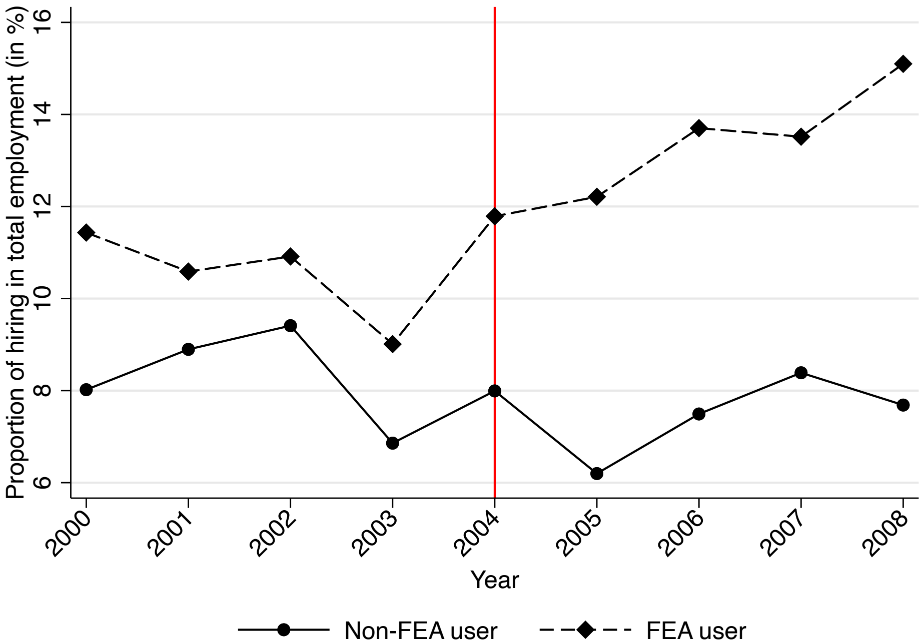
in which y_{it} represents the dependent variables 'share of hires' and 'employment growth' in establishment i at year t as calculated in Equations (1) and (2). The $HartzIII$ term is an indicator variable for the Hartz III reform and takes unit value after the reform was enacted on January 1st, 2004 and is zero otherwise. $FEAuser_i$ is an indicator variable for establishments using the Federal Employment Agency, which takes unit value in this case and is zero otherwise. Our difference-in-differences estimation strategy identifies the treatment effect on the treated (ATT) which is the treatment effect for those establishments using the agency relative to those establishments which do not (Imbens & Wooldridge 2009). This effect is identified by the coefficient τ of the interaction term in Equation (3). Establishment particularities in recruitment behavior are taken into account by industry fixed effects ρ_i and federal state fixed effects λ_i, which capture regional labor demand shocks at a given point in time. Since the Hartz reforms consist of three packages which are implemented successively we also add year fixed effects γ_t.[11] The idiosyncratic error term is denoted as ε_{it}.

[10]In the presence of heteroscedasticity, coefficient as well as standard error estimates in tobit models are inconsistent. It is, however, feasible to calculate a Wald test statistic to test for heteroscedasticity. Our applied Wald test clearly rejects the assumption of homoscedasticity and we therefore replace the variance σ with $\sigma_i = \sigma \times exp(w_i'\alpha)$ within the likelihood maximization (e.g. Greene 2008). The test statistic provides a value of 1199.19 with a p-value of .000. Hence, we apply a heteroscedastic tobit model in which we consider group-wise multiplicative heteroscedasticity. In this case α denotes estimated parameters of the heteroscedasticity term and w_i' is a vector of variables in which we include establishment-size as well as industry dummy variables to capture different hiring behavior among establishments and industries. Estimates from homoscedastic tobit models are also shown for reference.

[11]In a similar context of the Hartz reforms, Launov & Wälde (2016) for example capture other potential confounding reform effects using time dummy variables for the year 2002 and 2004.

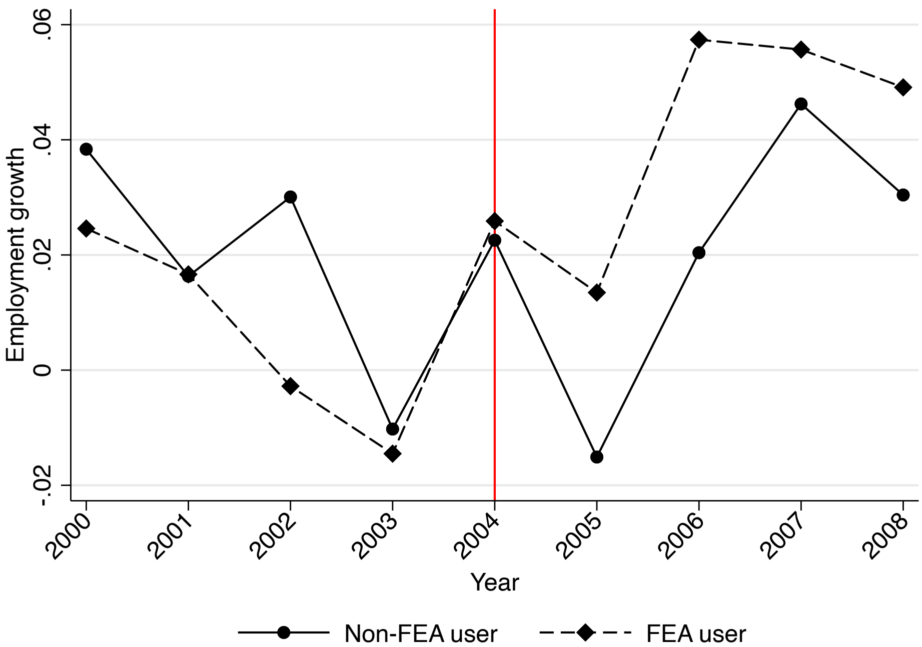
Control Variables. The vector X_m represents control variables in which we add a very comprehensive set of establishment and workforce characteristics. First, we use the logarithm of employees as well as the log of employees squared to account for establishment size effects since larger firms might tend to use the agency more frequently. We use an indicator variable identifying whether the firm is a stand-alone independent establishment or part of a firm group. This variable takes unit value when the establishment is not part of a firm group and is zero otherwise. To adjust for possible age effects of establishments in a sense that older establishments are more prone to use the employment agency, we include a dummy variable which takes unit value when the establishment was founded in the year 2000 or later and is zero otherwise. We take account of the possible influence of the legal form with the dummy variable 'limited liability'. Furthermore, we measure effects arising from coverage by a collective bargaining agreement with an indicator variable which takes unit value if the establishment is covered by a collective bargaining agreement and is zero otherwise. To control for different effects of concentrated ownership (one or few dominant owners) versus no dominant owner of the establishment, we use a dummy variable 'no dominant ownership' which has unit value if the ownership is broadly spread and is zero otherwise. Finally, to take employment expectations into account, we include a dummy variable which assumes unit value if the establishment indicates having such positive expectations and is zero otherwise. This variable is obviously relevant, because positive expectations will most likely result in plans for hiring and possibly the involvement of the Federal Employment Agency as well. Regarding the composition of the workforce, we include the share of part-time employees, female employees, highly qualified employees, fixed-term employees as well as the share of apprentices. Descriptive statistics which are differentiated according to the treatment and control groups are presented in Table 2.

Figure 2: Raw trend in share of hires in total employment



Notes: The figure shows the proportion of hiring in total employment as calculated in Equation (2) compared between the treatment and control group. Information on $N = 14,658$ observations on $N = 8,348$ treatment and $N = 6,310$ control observations. The treatment group consists of establishments that use the FEA and the control group of establishments that do not. The red line indicates the implementation of the Hartz III reform which became effective on January 1st, 2004. IAB Establishment Panel, waves 2000 to 2008.

Figure 3: Raw trend in employment growth



Notes: The figure shows the employment growth as calculated in Equation (1) compared between the treatment and control group. Information on $N = 14,658$ observations on $N = 8,348$ treatment and $N = 6,310$ control observations. The treatment group consists of establishments that use the Federal Employment Agency, while those in the control group do not. The red line indicates the implementation of the Hartz III reform which became effective on January 1st, 2004. IAB Establishment Panel, waves 2000 to 2008.

Table 2: **Descriptive statistics of treatment and control group (N=14,658)**

	Treatment group		**Control group**		**Differences**	
	$N = 8,348$		$N = 6,310$			
	Mean	Std. dev.	Mean	Std. dev.	(1)-(3)	p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variables						
Share of hires	12.16	29.96	7.96	20.08	4.19***	.000
Employment growth	.027	.223	.022	.223	.005	.202
Establishment controls						
(log) Employees	4.51	1.74	3.91	1.76	.601***	.000
(log) Employees squared	23.37	15.82	18.36	14.61	5.01***	.000
Western Germany	.718	.450	.771	.420	-.054***	.000
Founded in 2000 or later	.231	.421	.251	.433	-.020**	.004
Single Establishment	.608	.488	.634	.482	-.026***	.001
Limited Liability	.619	.486	.604	.489	.015*	.066
No dominant shareholder	.057	.231	.062	.242	-.006	.153
Positive employment expectations	.302	.459	.319	.466	-.017**	.028
Collective bargaining agreement	.523	.500	.473	.499	.050***	.000
Workforce controls						
Share of part time workers	.186	.226	.199	.241	-.013***	.001
Share of female workers	.410	.285	.414	.280	-.004	.356
Share of qualified workers	.691	.266	.699	.280	-.008*	.078
Share of fixed term workers	.094	.176	.055	.127	.038***	.000
Share of apprentices workers	.049	.087	.040	.075	.010***	.000

Notes: For the definition of the treatment and control group, see Section 4.2. Data from the IAB Establishment Panel, waves 2000–2008 with an overall sample size of $N = 14,658$. Column (5) displays the differences in mean values between treatment and control group and column (6) shows the corresponding p-value. Dependent variables are calculated as outlined in Section 4.3 in Equations (1) and (2). *, **, *** significant at the 10%, 5%, 1% level.

4.4 Baseline results and treatment effects

Baseline results on the Hartz III reform effect are considered in this section in which we estimate the difference-in-differences specification using ordinary least squares (OLS), fixed effects (FE) as well as tobit models as outlined in Section 4.3. The regression results for both dependent variables ('share of hires' and 'employment growth') are presented in Table 3.[12] In the case of the corner-solution tobit, marginal effects are presented. These marginal effects are computed at the intensive margin, which are the marginal effects for observations with values of the dependent variable above zero, which is $\mathbb{E}(Y|Y > 0)$ (McDonald & Moffitt 1980).

In our context, the most important variable in Equation (3) is the coefficient τ of the interaction term $HartzIII_t \times FEAuser_i$ which measures the impact of an increase in the efficiency of the Federal Employment Agency on the proportion of hires as well as employment growth in the treatment group compared to the control group. It turns out that the coefficient of this variable is positive and at least significant at the 5 percent level indicating a positive effect of the reform. In the baseline OLS models, the Hartz III reform increases the share of hires in the treatment group by roughly 2 percentage points compared to the control group. Evaluated at the sample mean of the share of hires variable, this corresponds to an increase by roughly 20% in which our results are in line with the findings by Launov & Wälde (2016), Krebs & Scheffel (2013) as well as Klinger & Rothe (2012). According to the tobit models, the effect is slightly smaller with point estimates ranging from 0.4 to 1.01 percentage points, which are, however, also significant. The Hartz III reform therefore seems to have a positive impact on the hiring rate in the treatment group relative to the control group.

The results on all variables included can be found in the Appendix Table A.2. First, our results show that younger establishments exhibit faster employment growth compared to older establishments which is quite in line with the literature (e.g. Haltiwanger et al. 2013). Variables capturing establishment size and age effects are highly significant. For

[12] See for the full specifications including all results for the control variables, Table A.2 in the Appendix.

Table 3: Results for OLS, FE and Tobit models

Dependent variable	Share of hires				Employment growth	
	OLS (1)	FE (2)	Tobit (3)	Het. Tobit (4)	OLS (5)	FE (6)
$HARTZIII$	-3.45***	-6.11***	-1.76***	-.832***	-.020**	-.093***
	(.852)	(1.30)	(.450)	(.152)	(.008)	(.015)
$FEAuser$	1.74***		1.05***	.041	-.014***	
	(.586)		(.304)	(.105)	(.005)	
$HARTZIII \times FEAuser$	2.01**	2.03**	.927**	.417***	.025***	.037***
	(.791)	(.864)	(.414)	(.141)	(.007)	(.014)
Establishment fixed effects		✓				✓
Industry fixed effects	✓	✓	✓	✓	✓	✓
Federal State fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓
R^2 / Pseudo R^2	.166	.861	.027	.034	.073	.820
Left (0) censored obs.			3,854	3,854		
Uncensored obs.			10,804	10,804		
Observations	14,658	14,658	14,658	14,658	14,658	14,658

Notes: IAB Establishment Panel, waves 2000–2008, with an overall sample size of $N = 14,658$ observations. Cluster-robust standard errors at the establishment level in parentheses. The control group consists of establishments that do not use the Federal Employment Agency, which is identified as reported in Section 4.2. The treatment group consists of establishments that use the placement services. The latter group, therefore, is affected by the Hartz III reform which was implemented on January 1st, 2004. Estimation regarding the specification Equation (3). Tobit model denotes the homoscedastic tobit model and, in the heteroscedastic tobit model, we include a vector of establishment size and industry dummy variables for the variance estimation. For more information regarding the heteroskedastic tobit model see Section 4.3. Year fixed effects include year dummy variables ranging from the year 2001 to 2008 with the year 2000 being the base category. Control variables are included as outlined in Section 4.3. Fixed effects are nested within establishment cluster. *, **, *** significant at the 10%, 5%, 1% level.

single establishments not belonging to a multi-plant firm, we only find significant effects for the tobit specification and the employment growth variable. These positive effects may arise in particular because of replacement hires for employees who have left the establishment. More interestingly, though, are the effects of the workforce composition. First, we observe that the coefficient of fixed-term contracts is highly significant and positive indicating that establishments might rely on a large share of fixed-term workers to fill vacancies. Second, a higher share of apprentices seems to decrease the share of hires. It might be the case that workers stay in the job after their apprenticeship, which decreases the need for new recruitments. Finally, we also find significant effects for the limited liability coefficient, which is positive for the share of hires but negative for the employment growth variable. Thus, limited liability establishments may therefore have a higher fraction of hires, although they may grow at a lower pace compared to other establishments which do not the legal form of a limited liability.

4.5 Structural break test for the Hartz III reform

In this section we want to check whether the effect of the Hartz III reform can be measured not only by the treatment effect coefficient τ in Equation (3), but also by changes in all covariates. We therefore draw on the structural break literature (e.g. Chow 1960; Gujarati 1970; Dufour 1982; Cantrell et al. 1991; Antoch et al. 2019) to support our difference-in-differences results from Table 3. The literature on change point detection is well developed and, besides studies in a time series context, recent empirical applications in particular also consider the panel data context (e.g. Jayachandran et al. 2010; Wiese 2014; Antoch et al. 2019; Lunsford 2020).

With (a variant of) the Chow test, we investigate whether the Hartz III reform may well have changed the effects of a large number of variables. We therefore test whether the Hartz III reform not only constitutes a shift in our dependent variables as shown in Table 3, but also affects the whole set of control variables as well. In this view, the Hartz III reform constitutes a regime shift in terms of recruitment behavior in which the reform also affects other establishment characteristics. To do so, we apply a more generalized version of the Chow test (Chow 1960; Cantrell et al. 1991) using the dummy variable technique as proposed by Gujarati (1970) which is for example applied in Lunsford (2020). Whereas the classical Chow test provides evidence for the difference between two regression models, the dummy variable approach is also able to specify the source of difference, i.e. which is either due to the intercept, the slope or both (Smith 2015; Gujarati 1970).

To implement this approach we augment our baseline specification (presented in Section 4.3) by adding a set of interaction variables consisting of the control variables multiplied by Hartz III dummy variable. This specification is shown in Equation (4). We expect a significant break point at the timing of the Hartz III reform in the year 2004 in the series for the treatment group, but not in the control group since the latter group is unaffected by the reform.

We perform the estimates for Equation (4) separately for the treatment and control observations for our two dependent variables y_{it}, i.e. the 'share of hires' and 'employment growth'.[13] Consider the following regressions which we separately fit for the treatment and control group, which are denoted as $g = (T, C)$.

$$y_{itg} = \alpha_{1g} + \alpha_{2g} HartzIII_{tg} + \beta_{1g} X_{mg} + \beta_{2g} HartzIII_{tg} \times X_{mg} + \varepsilon_{itg} \tag{4}$$

where $i = 1, \ldots, N$ are the observations within the treatment and control group. $t = 2000, \ldots, 2008$ and the indicator variable $HartzIII_{tg}$ is defined as $HartzIII = 0$ if the year equals 2000–2003 and $HartzIII = 1$ if the year is equal to 2004–2008. In using the generalized dummy variable Chow approach (Gujarati 1970; Cantrell et al. 1991; Lunsford 2020), we inspect the following sources of structural change due to the Hartz III reform:

$$\begin{aligned} \mathbb{E}(y_{it}|HartzIII = 0) &= \alpha_{1g} + \beta_{1g} X_{mg} \\ \mathbb{E}(y_{it}|HartzIII = 1) &= \underbrace{(\alpha_{1g} + \alpha_{2g})}_{\text{Break in intercept}} + \underbrace{(\beta_{1g} + \beta_{2g})}_{\text{Break in slope}} X_{mg} \end{aligned} \tag{5}$$

For each of our two dependent variables y_{itg} and for the treatment and control group $g = (T, C)$ we perform Wald-tests on the α and β coefficients separately to check whether the α or β coefficients are jointly different from zero to test for a structural break in the intercept or the slope. We therefore perform eight different regressions and, if the reform effects are strong enough, we should see a significant difference in the treatment group but not in the control group. Results of these tests are provided in Table 4.

As expected, and shown in Table 4, the test results indicate no reform effect on the establishments forming the control group, neither for the employment growth variable nor for the share of hires. For observations from the treatment group, however, there are significant differences between the pre- and post-intervention Hartz III period. We therefore find supplementary evidence besides the difference-in-differences estimation, that there is indeed a reform effect in the Hartz III affected treatment group, but not in the control group. Furthermore, the dummy variable approach Chow test (Gujarati 1970;

[13] The industry as well as federal state fixed effects which are denoted ρ_i in Equation (3) are in this specification summarized within the X_m control variables. They are also subject to a potential break point.

Table 4: **Results for structural break tests**

	Employment growth		**Share of hires**	
Break point:	Treatment	Control	Treatment	Control
Hartz III (2004)	(1)	(2)	(3)	(4)
Break in intercept	1.84	1.95	.610	2.40
	(.175)	(.163)	(.435)	(.121)
Break in slope	2.09***	1.23	1.52***	1.04
	(.000)	(.120)	(.004)	(.398)
Industry fixed effects	✓	✓	✓	✓
Federal state fixed effects	✓	✓	✓	✓
Control variables	✓	✓	✓	✓
R^2	.086	.090	.211	.125
Observations	8,348	6,310	8,348	6,310

Notes: This table shows the dummy variable technique Chow test according to the specification in Equations (4) and (5) as outlined in Gujarati (1970). Results show the Wald-statistic and the corresponding p-value in parenthesis. The Wald test is calculated for both dependent variables between the treatment and control group including the industry and federal state dummy variables. Critical values for the test statistics differ because of different sample sizes and thus degrees of freedom between the samples. Control variables are included as outlined in Section 4.3. Results account for selection effects using IPW weights as outlined in Section 5.1. Different weights provide very similar results. Significance: *, **, *** significant at the 10%, 5%, 1% level.

Cantrell et al. 1991; Lunsford 2020) allows us to test whether the structural break arises because of shifts in the intercept or the slope coefficients. For the 'employment growth' and 'share of hires' variables, we find significant differences for the slope coefficients but not for the intercepts.

5 Robustness

5.1 Selectivity of Federal Employment Agency usage

Although our difference-in-differences model includes a large set of establishment control variables, there might also be pre-existing differences determining the FEA user status which is not captured by these variables. For example, establishments might need highly specialized personnel for whom the employment office is not the right service provider. Another possibility is an unobserved demand shock, which has a positive effect on the growth opportunities and simultaneously causes companies to contact the employment office, having not had to do so previously. If this is actually the case, we face a selection problem since unobserved variables affecting both the decision to use the employment agency as well determinants of employment growth.

We tackle this problem by applying difference-in-differences estimation with an inverse probability weighing (IPW) approach (e.g. Imbens & Wooldridge 2009).[14] The idea behind this approach is to create a similar sample of establishments in which the treatment (FEA usage) is independent of observed confounders. This process follows a two-step approach. First, we use the binary dependent variable which is defined as the treatment indicator and takes unit value if the establishment uses the employment agency for their recruitment and zero otherwise. Then, we estimate the propensity score p_t for each available year from 2000 to 2008 using a probit model for binary dependent variables. We adjust for the composition of the workforce by including the share of part-time workers, female workers, highly qualified workers, apprentices and the share of workers employed on the basis of fixed-term contracts. We also include a comprehensive set of control variables which are the same as in the regressions in Equations (3) and (6). We also take industry fixed effects into account. The results of the probit regressions used to calculate the propensity score for each year are presented in Table A.3 of the Appendix. Second, we calculate the inverse of these obtained propensity scores to re-weight the difference-in-differences regressions accordingly.[15] Finally, we provide mean comparisons between the FEA users and non-users which are provided in Table A.4 of the Appendix.

As shown in the last column of Table A.4, all differences in covariates between the treatment and control group are vanished after the IPW matching procedure. Results of these re-weighted regressions are presented in Table 5. As before, the interaction term denotes the treatment effect, which is positive and significant for the OLS, fixed effects

[14]For a similar approach in the context of unemployment benefits and re-employment rates, see for example Uusitalo & Verho (2010). In Section 5.3 we also apply different definitions of the IPW approach in which we additionally use propensity score trimming, different weights as well as normalized weights.

[15]The control group then receives the weights which are calculated as $w_t^c = \frac{1}{(1-p_t)}$ and the treatment group receives weights which are calculated as $w_t^t = \frac{1}{p_t}$ (e.g. Imbens & Wooldridge 2009). For different specifications of the weights, see Section 5.3.

Table 5: Results of IPW OLS, FE and Tobit models

Dependent variable	Share of hires				Employment growth	
	OLS (1)	FE (2)	Tobit (3)	Het. Tobit (4)	OLS (5)	FE (6)
$HARTZIII$	-3.66 ***	-5.64***	-1.72***	-.836***	-.025***	-0.79***
	(.920)	(1.17)	(.472)	(.172)	(.009)	(.014)
$FEAuser$	1.99***		1.14***	.042	-.013**	
	(.565)		(.293)	(.110)	(.005)	
$HARTZIII \times FEAuser$	1.83**	1.92**	.704*	.313*	.025***	.029**
	(.804)	(.931)	(.422)	(.170)	(.008)	(.013)
Establishment fixed effects		✓				✓
Industry fixed effects	✓	✓	✓	✓	✓	✓
Federal state fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓
R^2 / Pseudo R^2	.155	.875	.028	.037	.074	.837
Left (0) censored obs.			3,847	3,847		
Uncensored obs.			10,770	10,770		
Observations	14,617	14,617	14,617	14,617	14,617	14,617

Notes: IAB Establishment Panel, waves 2000–2008. Cluster-robust standard errors at the establishment level in parentheses. Estimation regarding the specification in Equation (3). Tobit model denotes the homoscedastic tobit model and in the heteroskedastic tobit model we include a vector of establishment size and industry dummy variables for the variance estimation. Heteroscedastic Tobit specification as in Section 4.3. Year fixed effects include year dummy variables ranging from the year 2001 to 2008 with the year 2000 being the base category. Control variables are included as outlined in Section 4.3. The control group receives the weights which are calculated as $w_t^c = \frac{1}{(1-p_t)}$ and the treatment group receive weights which are calculated as $w_t^t = \frac{1}{p_t}$. Here, p_t is the propensity score for each cross-section calculated as the predicted probability of receiving the treatment stemming from probit estimates in Table A.3. For robustness tests regarding the calculation of weights see Section 5.3. Fixed effects are nested within establishment cluster. Significance: *, **, *** significant at the 10%, 5%, 1% level.

and tobit specifications. The coefficients in the difference-in-differences regressions are very similar to the unweighted ones, presented in Table 3. After the Hartz III reform is in place, establishments using the Federal Employment Agency have a 1.8-percentage-point increased share of hires compared to the establishments not using the placement services. These results are still significant at the 1 percent level. The marginal effects based on the tobit estimations are again smaller than the OLS coefficients and also slightly smaller than the marginal effects presented in Table 3 but they remain significant. In terms of employment growth, our results show that establishments that use the placement services indeed also have a higher employment growth in the magnitude of 2.5 percentage points.

5.2 Test for common trend before the Hartz III reform

A crucial assumption for the identification of the treatment effect within the difference-in-differences framework is the common trend assumption. It states that trends in outcome variables among the treatment and control group should be similar before the intervention (e.g. Imbens & Wooldridge 2009). In our case this assumption states that the treatment group have the same trend in employment growth and share of hires before the Hartz III intervention. As for example shown in Figure 3 and 2, the unadjusted raw trends in both dependent variables are roughly similar before the intervention. After the reform was implemented, however, both trends diverge.

To test the common trend assumption, we apply the following augmented regression for both dependent variables (e.g. Mora & Reggio 2015). To do so, we re-estimate the model given in Equation (3) and replace the Hartz III dummy variable and interaction with a set of time dummies and its interaction terms with the treatment dummy, resulting in the model presented in Equation (6). A similar approach in this context is also provided for example by Giebel & Kraft (2019) and Hangoma et al. (2018).

$$\begin{aligned} y_{it} = \alpha + \beta_1 FEAuser_i + \sum_{t=2001}^{2008} \tau_t \times FEAuser_i \times Year_t \\ + \beta_m X_m + \gamma_t + \rho_i + \varepsilon_{it} \end{aligned} \tag{6}$$

In this setting, y_{it} are the dependent variables as outlined in Section 4.3. X_m is a vector of control variables. A set of industry fixed effects is denoted as ρ_i, year fixed effects as γ_t. and the idiosyncratic error term is denoted as ε_{it}. The estimation results which also include control variables are presented in Table 6.[16]

For the common-trend to hold, we test whether all year FEA user interaction variables in the pre-treatment period before the year 2004 are jointly not different from zero. Thus, we test the parallel trend assumption with $H_0 : \tau_t = 0 \; \forall \; t \leq 2003$. By estimating Equation

[16] We also estimated fixed effects models using this specification: However, the results of the test statistics do not change much and we do not reject H_0.

Table 6: Flexible model and test for common trends

Dependent variable	Share of hires		Employment growth	
	Het. Tobit	Het. Tobit IPW	OLS	OLS IPW
	(1)	(2)	(3)	(4)
FEA user	-.009	-.001	-.014	-.013
	(.414)	(.450)	(.009)	(.010)
FEA user × 2001	.529	.356	.014	.014
	(.506)	(.601)	(.013)	(.014)
FEA user × 2002	.067	-.001	-.019	-.020
	(.561)	(.740)	(.014)	(.014)
FEA user × 2003	-.553	-1.069	.003	.004
	(.752)	(.740)	(.016)	(.017)
FEA user × 2004	.816	1.456*	.008	.011
	(.631)	(.834)	(.016)	(.018)
FEA user × 2005	.655	-.860	.035**	.044***
	(.602)	(.994)	(.016)	(.019)
FEA user × 2006	1.515**	1.144*	.048***	.041**
	(.604)	(.692)	(.016)	(.017)
FEA user × 2007	.972*	.323	.015	.009
	(.577)	(.601)	(.014)	(.015)
FEA user × 2008	1.312**	1.225*	.021	.026*
	(.551)	(.666)	(.013)	(.015)
Constant	-11.96***	-10.09***	-.167***	-.171***
	(4.10)	(3.83)	(.028)	(.029)
$H_0 : \tau_t = 0 \; \forall \; t \leq 2003$:	2.62	3.16	1.86	1.53
F / Wald-statistic (p-value)	(.455)	(.368)	(.135)	(.205)
Industry fixed effects	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓
Control variables	✓	✓	✓	✓
R^2/ Pseudo R^2	.164	.150	.074	.075
Observations	14,658	14,617	14,658	14,617

Notes: IAB Establishment Panel, waves 2000–2008 with an overall sample size of $N = 14,658$ observations and $N = 14,617$ observations for the IPW re-weighted estimation results in column (2) and (4). Cluster-robust standard errors at the establishment level in parentheses. Estimation regarding the specification in Equation (6) in which the treatment effect is shown over time. Stated null hypothesis tests for common pre-treatment trends (i.e. joint significance of treatment-year interaction terms before the year 2004.) More details are provided in Section 5. Point estimates and test results are also more or less the same if we apply different weighting schemes as explained in the next Section 5.3. Significance: *, **, *** significant at the 10%, 5%, 1% level.

(6) we test for the joint significance of the pre-treatment year-treat interaction effects. As before we use OLS to explain employment growth and (heteroscedastic) tobit to explain hire rates. Furthermore we also estimated the selectivity-adjusted IPW models explained and presented in Section 5.1.[17] For the share of hires in column (1) we perform Wald-

[17] We also estimated these models using inverse probability reweighed models in which we apply different weights, as well as propensity score trimming and normalized weights as outlined in Section 5.3. The

tests and find $Wald = 2.62$ with a p-value of .455; and (2) for the share of hires (IPW) weighted: $Wald = 3.16$ with a p-value of .368. For the employment growth dependent variable, we obtain in column (3) unweighted: $F = 1.86$ with a p-value of .135 and (4) (IPW) weighted: $F = 1.53$ with a p-value of .205. The joint F-tests reveal that indeed we cannot reject the possibility that all pre-treatment year-treat interaction effects are different from zero. Thus, the common trend assumption seems to be fulfilled. To sum up, the results of the estimation of Equation (6) supports the common trend assumption in which the trends of the employment growth and share of hires are equal before the Hartz III intervention.

5.3 Different IPW weights

Calculating the average treatment effect on the treated (ATT). The results so far have to be interpreted as the average treatment effect (ATE) in which we use calculated weights as described in Section 5.1. As a further robustness test, we calculate different weights and calculate the average treatment effect on the treated (ATT) (e.g. Stuart 2010). Compared to the ATE results in which we weight the treatment and control group, we now only re-weight the comparison group to match the distribution of control variables compared to the treatment group. Thus, the control group receives weights which are calculated as $w_t^c = \frac{p_t}{(1-p_t)}$ and the treatment group receives weights $w_t^t = 1$. Similarly as before, p_t is the propensity score for each cross-section calculated as the predicted probability of using the Federal Employment Agency stemming again from probit estimates provided in Table A.3 of the Appendix.[18] Results are provided in Table 7.

Reweighted estimates using ATT weights are usually slightly larger in magnitudes (e.g. Uusitalo & Verho 2010), which is what we also find in our results for most of the specifications. The results are, however, at similar levels of significance. In all specifications using the selectivity adjusted difference-in-differences specification we find robust and sig-

results of the common trend tests, however, did not change and there is no specification in which we reject H_0. We thus conclude that the common trends assumptions is met in our sample.

[18]See for a similar application (Campolieti 2018; Uusitalo & Verho 2010).

Table 7: **Results of IPW–OLS, FE and Tobit models using ATT weights**

Dependent variable	Share of hires				Employment growth	
	OLS (1)	FE (2)	Tobit (3)	Het. Tobit (4)	OLS (5)	FE (6)
$HARTZIII$	-3.88***	-5.78***	-1.74***	-.922***	-.024**	-0.75***
	(1.09)	(1.38)	(.542)	(.212)	(.010)	(.014)
$FEAuser$	1.98***		1.12***	-.010	-.012**	
	(.593)		(.304)	(.131)	(.006)	
$HARTZIII \times FEAuser$	2.72***	1.79*	1.21***	.514**	.028***	.025*
	(.888)	(1.02)	(.455)	(.209)	(.008)	(.013)
Establishment fixed effects		✓				✓
Industry fixed effects	✓	✓	✓	✓	✓	✓
Federal state fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓
R^2 / Pseudo R^2	.177	.870	.027	.039	.067	.827
Left (0) censored obs.			3,847	3,847		
Uncensored obs.			10,770	10,770		
Observations	14,617	14,617	14,617	14,617	14,617	14,617

Notes: IAB Establishment Panel, waves 2000–2008. Cluster-robust standard errors at the establishment level in parentheses. Estimation regarding the specification in Equation (3). Year fixed effects include year dummy variables ranging from the year 2001 to 2008 with the year 2000 being the base category. Control variables are included as outlined in Section 4.3. The control group receives the weights which are calculated as $w_t^c = \frac{p_t}{(1-p_t)}$ and the treatment group receive weights which are calculated as $w_t^t = 1$. Here, p_t is the propensity score for each cross-section calculated as the predicted probability of receiving the treatment stemming from probit estimates in Table A.3. For robustness tests regarding the calculation of weights see Section 5.3. Fixed effects are nested within establishment cluster. Significance: *, **, *** significant at the 10%, 5%, 1% level.

nificant positive employment effects for the FEA user group compared to the non-user group after the reform was in place. Comparisons of means among the covariates after the re-weighting approach are also balanced which is a necessary condition for interpreting the results. See Table A.7 for the balancing of covariates.

IPW Trimming and normalized weights. Extreme values of the weights might impose a threat to the identification of the treatment effect and the variance of the estimates (e.g. Kranker et al. 2020). This rationale holds for both the ATE and the ATT results. The usual solution to this threat relies on dropping values with extreme large or small weights. As a further robustness check we therefore apply symmetric trimming in which we exclude establishments with propensity scores outside of the range of $[\alpha, 1-\alpha]$ in which α is a threshold parameter which can be chosen by the researcher (Li et al. 2018). We choose a quite common value of $\alpha = 0.1$ and discard those establishments with propensity scores below and above the threshold to ensure a better overlap (Crump et al.

2009). We therefore lose 421 observations for the following regressions.

As a final step, we normalize weights to sum to one when estimating the reweighted difference-in-differences specifications (Busso et al. 2014). There are in fact many empirical examples in which normalized matching estimators are used in the empirical literature (e.g. Robins et al. 2007; Imbens 2004). They provide some efficiency advantages and moreover, they are more reliable in finite samples (e.g. Busso et al. 2014). Results for using trimmed and normalized ATT weights are provided in Table 8 and results for ATE weights in Table 9. As a final robustness check we also estimated each specification using either (i) only propensity score trimming with a similar trimming value of $\alpha = 0.1$ or (ii) normalized weights. For every specification in which we estimate OLS, fixed effects as well as tobit models, we find very similar results as presented in the Tables 8 and 9.

Table 8: Trimmed propensity score and normalized ATT weights

Dependent variable	Share of hires				Employment growth	
	OLS (1)	FE (2)	Tobit (3)	Het. Tobit (4)	OLS (5)	FE (6)
$HARTZIII$	-2.94***	-4.37**	-1.39***	-.763***	-.016*	-.075***
	(.993)	(1.09)	(.500)	(.183)	(.009)	(.014)
$FEAuser$	2.29***		1.26***	.066	-.010*	
	(.580)		(.297)	(.127)	(.005)	
$HARTZIII \times FEAuser$	2.05**	1.97**	.941**	.442***	.019**	.023*
	(.843)	(.953)	(.434)	(.171)	(.008)	(.013)
Establishment fixed effects		✓				✓
Industry fixed effects	✓	✓	✓	✓	✓	✓
Federal state fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓
R^2 / Pseudo R^2	.158	.893	.028	.036	.066	.833
Left (0) censored obs.			3,757	3,757		
Uncensored obs.			10,439	10,439		
Observations	14,196	14,196	14,196	14,196	14,196	14,196

Notes: IAB Establishment Panel, waves 2000–2008. Cluster-robust standard errors at the establishment level in parentheses. Year fixed effects include year dummy variables ranging from the year 2001 to 2008 with the year 2000 being the base category. Control variables are included as outlined in Section 4.3. The control group receives the weights which are calculated as $w_t^c = \frac{p_t}{(1-p_t)}$ and the treatment group receives weights which are calculated as $w_t^t = 1$. We also apply symmetric trimming using the threshold parameter $\alpha = 0.1$ as well as normalized IPW weights (e.g. Busso et al. 2014). Fixed effects are nested within establishment cluster. Significance: *, **, *** significant at the 10%, 5%, 1% level.

Table 9: Trimmed propensity score and normalized ATE weights

Dependent variable	Share of hires				Employment growth	
	OLS (1)	FE (2)	Tobit (3)	Het. Tobit (4)	OLS (5)	FE (6)
$HARTZIII$	-3.04***	-4.77***	-1.45***	-.740***	-.018**	-.080***
	(.885)	(1.05)	(.457)	(.156)	(.009)	(.014)
$FEAuser$	2.21***		1.23***	.089	-.011**	
	(.563)		(.292)	(.108)	(.005)	
$HARTZIII \times FEAuser$	1.46*	2.07**	.606*	.310**	.019**	.030**
	(.791)	(.929)	(.353)	(.145)	(.008)	(.013)
Establishment fixed effects		✓				✓
Industry fixed effects	✓	✓	✓	✓	✓	✓
Federal state fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓
R^2 / Pseudo R^2	.142	.890	.029	.033	.070	.840
Left (0) censored obs.			3,757	3,757		
Uncensored obs.			10,439	10,439		
Observations	14,196	14,196	14,196	14,196	14,196	14,196

Notes: IAB Establishment Panel, waves 2000–2008. Cluster-robust standard errors at the establishment level in parentheses. Year fixed effects include year dummy variables ranging from the year 2001 to 2008 with the year 2000 being the base category. Control variables are included as outlined in Section 4.3. The control group receives the weights which are calculated as $w_t^c = \frac{1}{(1-p_t)}$ and the treatment group receives weights which are calculated as $w_t^t = \frac{1}{p_t}$. We also apply symmetric trimming using the threshold parameter $\alpha = 0.1$ as well as normalized IPW weights (e.g. Busso et al. 2014). Fixed effects are nested within establishment cluster. Significance: *, **, *** significant at the 10%, 5%, 1% level.

6 Conclusion

Since their introduction, the Hartz reforms have been the subject of much controversy, and the intensity of this discussion is increasing rather than decreasing. Our contribution focuses on a less acknowledged part of the reforms, namely the modernization of the employment agency stipulated in the Hartz III reform. In this paper we analyze an increase in the job placement efficiency of the Federal Employment Agency on employment growth. Compared to other studies, we measure the effect not on the individual or macro level, but rather on the establishment level.

A unique exogenous shock arising from the Hartz III legislation in the matching technology of the agency in Germany in 2004 allows us to investigate hiring behavior and employment growth on the labor demand side. We use the IAB Establishment Panel provided by the Institute for Employment Research (IAB) to identify establishments that actually use the placement services of the Federal Employment Agency and compare those to the control group of firms which do not. We apply conditional difference-in-differences estimations to measure the treatment effect on the treated ones. In addition, we take selectivity issues for the decision to use the placement service into account by applying inverse probability weighting. We provide evidence that the reform, which re-framed the agency, is indeed beneficial for job placement. Our estimates show that establishments which use the services achieve an increase in the proportion of hires, and the employment growth is also higher compared to establishments which do not use the placement services. These results are robust to selectivity, which we checked using inverse-probability weighting with different specifications for the weights. The common trend assumption also seems to be fulfilled prior to the Hartz III intervention.

Our paper, however, is not without limitations. An important extension to our study is the differentiation of employment for example between temporary and permanent employment. The Federal Employment Agency may be particularly relevant for unskilled and low-educated workers (e.g. Fougère et al. 2009), thus, there may be substitution effects in a sense that firms substitute costly permanent employment with temporary agency workers.

With respect to policy implications, we provide further evidence for the importance of the placement service in the labor market. The need for efficient placement agencies will probably increase even more if, for example, members of certain qualification groups (low but also medium-qualified) are dismissed because of technical progress. Getting them back into work requires efficient matching. The current problems on the labor market in the context of the COVID-19 crisis will also result in an additional need for efficient job searching. A modernization of public institutions, as in Germany via the Hartz III reform, can be helpful in this context.

References

Acemoglu, D. & Restrepo, P. (2018). Modeling Automation. *AEA Papers and Proceedings*, *108*, 48–53.

Akyol, M., Neugart, M., & Pichler, S. (2013). Were the Hartz Reforms Responsible for the Improved Performance of the German Labour Market? *Economic Affairs*, *33*(1), 34–47.

Antoch, J., Hanousek, J., Horváth, L., Hušková, M., & Wang, S. (2019). Structural breaks in panel data: Large number of panels and short length time series. *Econometric Reviews*, *38*(7), 828–855.

Arent, S. & Nagl, W. (2013). Unemployment Compensation and Wages: Evidence from the German Hartz Reforms. *Jahrbücher für Nationalökonomie und Statistik*, *233*(4), 450–466.

Autor, D. H. (2001). Wiring the Labor Market. *Journal of Economic Perspectives*, *15*(1), 25–40.

Battisti, M., Giesing, Y., & Laurentsyeva, N. (2019). Can job search assistance improve the labour market integration of refugees? Evidence from a field experiment. *Labour Economics*, *61*, 101745.

Bauer, A. & King, I. (2018). The Hartz reforms, the German Miracle, and labor reallocation. *European Economic Review*, *103*, 1–17.

Blasco, S. & Pertold-Gebicka, B. (2013). Employment policies, hiring practices and firm performance. *Labour Economics*, *25*, 12–24.

Bossler, M., Geis, G., & Stegmaier, J. (2018). Comparing survey data with an official administrative population: assessing sample-selectivity in the IAB Establishment Panel. *Quality and Quantity*, *52*(2), 899–920.

Bradley, J. & Kügler, A. (2019). Labor market reforms: An evaluation of the Hartz policies in Germany. *European Economic Review*, *113*, 108–135.

Brändle, T. & Goerke, L. (2018). The One Constant: A Causal Effect of Collective Bargaining on Employment Growth? *Scottish Journal of Political Economy*, *65*(5), 445–478.

Brynjolfsson, E. & McAfee, A. (2014). *The second machine age work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Bryson, A. (2004). Managerial responsiveness to union and nonunion worker voice in Britain. *Industrial Relations*, *43*(1), 213–241.

Bryson, A. & Nurmi, S. (2011). Private sector employment growth, 1998-2004: a panel analysis of British workplaces. *Cambridge Journal of Economics*, *35*(1), 85–104.

Busso, M., DiNardo, J., & McCrary, J. (2014). New Evidence on the Finite Sample Properties of Propensity Score Reweighting and Matching Estimators. *The Review of Economics and Statistics*, *96*(5), 885–897.

Campolieti, M. (2018). Matching and Inverse Propensity Weighting Estimates of the Union Wage Premium: Evidence from Canada, 1997–2014. *Industrial Relations: A Journal of Economy and Society*, *57*(1), 101–130.

Cantrell, R. S., Burrows, P. M., & Vuong, Q. H. (1991). Interpretation and Use of Generalized Chow Tests. *International Economic Review*, *32*(3), 725–741.

Chodorow-Reich, G. (2014). The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008–9 Financial Crisis. *The Quarterly Journal of Economics*, *129*(1), 1–59.

Chow, G. C. (1960). Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica*, *28*(3), 591–605.

Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, *96*(1), 187–199.

Davis, S. J., Faberman, R. J., & Haltiwanger, J. C. (2013). The Establishment-Level Behavior of Vacancies and Hiring. *The Quarterly Journal of Economics*, *128*(2), 581–622.

Davis, S. J. & Haltiwanger, J. (1992). Gross Job Creation, Gross Job Destruction, and Employment Reallocation. *The Quarterly Journal of Economics*, *107*(3), 819–863.

Dlugosz, S., Stephan, G., & Wilke, R. A. (2014). Fixing the leak: Unemployment incidence before and after a major reform of unemployment benefits in Germany. *German Economic Review*, *15*(3), 329–352.

Dufour, J.-M. (1982). Generalized Chow Tests for Structural Change: A Coordinate-Free Approach. *International Economic Review*, *23*(3), 565–575.

Dustmann, C., Fitzenberger, B., Schönberg, U., & Spitz-Oener, A. (2014). From Sick Man of Europe to Economic Superstar: Germany's Resurgent Economy. *Journal of Economic Perspectives*, *28*(1), 167–188.

Ehrich, M., Munasib, A., & Roy, D. (2018). The Hartz reforms and the German labor force. *European Journal of Political Economy*, *55*, 284–300.

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.

Fahr, R. & Sunde, U. (2009). Did the Hartz Reforms Speed-Up the Matching Process? A Macro-Evaluation Using Empirical Matching Functions. *German Economic Review*, *10*(3), 284–316.

Fischer, G., Janik, F., Müller, D., & Schmucker, A. (2009). The IAB Establishment Panel – Things Users Should Know. *Schmollers Jahrbuch*, *129*(1), 133–148.

Fleckenstein, T. (2008). Restructuring welfare for the unemployed: The Hartz legislation in Germany. *Journal of European Social Policy*, *18*(2), 177–188.

Fougère, D., Pradel, J., & Roger, M. (2009). Does the public employment service affect search effort and outcomes? *European Economic Review*, *53*(7), 846–869.

Frey, C. B. & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, *114*, 254–280.

Gehrke, B., Lechthaler, W., & Merkl, C. (2019). The German labor market during the Great Recession: Shocks and institutions. *Economic Modelling*, *78*, 192–208.

Giebel, M. & Kraft, K. (2019). External Financing Constraints and Firm Innovation. *The Journal of Industrial Economics*, *67*(1), 91–126.

Gralla, R. & Kraft, K. (2018). Separating Introduction Effects from Selectivity Effects: The Differences in Employment Patterns of Codetermined Firms. *LABOUR*, *32*(1), 93–111.

Greene, W. H. (2008). *Econometric Analysis*. Upper Saddle River, NJ: Prentice Hall.

Gujarati, D. (1970). Use of Dummy Variables in Testing for Equality between Sets of Coefficients in Two Linear Regressions: A Note. *The American Statistician*, *24*(1), 50–52.

Haltiwanger, J., Jarmin, R. S., & Miranda, J. (2013). Who Creates Jobs? Small versus Large versus Young. *Review of Economics and Statistics*, *95*(2), 347–361.

Hangoma, P., Robberstad, B., & Aakvik, A. (2018). Does Free Public Health Care Increase Utilization and Reduce Spending? Heterogeneity and Long-Term Effects. *World Development*, *101*, 334–350.

Hartung, B., Jung, P., & Kuhn, M. (2018). What Hides behind the German Labor Market Miracle? Unemployment Insurance Reforms and Labor Market Dynamics. *CEPR Discussion Paper No. DP13328*.

Holzner, C. & Watanabe, M. (2015). Understanding the Role of Public Employment Services: Theory and Empirical Evidence. *IZA Conference Paper, Bonn*, 32.

Horton, J. J. (2017). The Effects of Algorithmic Labor Market Recommendations: Evidence from a Field Experiment. *Journal of Labor Economics*, *35*(2), 345–385.

Hud, M. & Hussinger, K. (2015). The impact of R&D subsidies during the crisis. *Research Policy*, *44*(10), 1844–1855.

Imbens, G. W. (2004). Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *The Review of Economics and Statistics*, *86*(1), 4–29.

Imbens, G. W. & Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, *47*(1), 5–86.

Jacobi, L. & Kluve, J. (2007). Before and after the Hartz reforms: The performance of active labour market policy in Germany. *Journal for Labour Market Research*, *40*(1), 45–64.

Jayachandran, S., Lleras-Muney, A., & Smith, K. V. (2010). Modern medicine and the twentieth century decline in mortality: Evidence on the impact of sulfa drugs. *American Economic Journal: Applied Economics*, *2*(2), 118–146.

Kaas, L. & Kircher, P. (2015). Efficient Firm Dynamics in a Frictional Labor Market. *American Economic Review*, *105*(10), 3030–3060.

Klinger, S. & Rothe, T. (2012). The Impact of Labour Market Reforms and Economic Performance on the Matching of the Short-term and the Long-term Unemployed. *Scottish Journal of Political Economy*, *59*(1), 90–114.

Klinger, S. & Weber, E. (2016). Decomposing Beveridge Curve Dynamics By Correlated Unobserved Components. *Oxford Bulletin of Economics and Statistics*, *78*(6), 877–894.

Kranker, K., Blue, L., & Forrow, L. V. (2020). Improving Effect Estimates by Limiting the Variability in Inverse Propensity Score Weights. *The American Statistician*, *0*(0), 1–12.

Krebs, T. & Scheffel, M. (2013). Macroeconomic evaluation of labor market reform in Germany. *IMF Economic Review*, *61*(4), 664–701.

Launov, A. & Wälde, K. (2016). The employment effect of reforming a public employment agency. *European Economic Review*, *84*, 140–164.

Le Barbanchon, T. (2016). The effect of the potential duration of unemployment benefits on unemployment exits to work and match quality in France. *Labour Economics*, *42*, 16–29.

Li, F., Thomas, L. E., & Li, F. (2018). Addressing Extreme Propensity Scores via the Overlap Weights. *American Journal of Epidemiology*, *188*(1), 250–257.

Liechti, F. (2020). Connecting Employers and Workers: Can Recommendations from the Public Employment Service Act as a Substitute for Social Contacts? *Work, Employment and Society*, *34*(4), 587–604.

Lunsford, K. G. (2020). Policy Language and Information Effects in the Early Days of Federal Reserve Forward Guidance. *The American Economic Review*, *110*(9), 2899–2934.

Marinescu, I. & Rathelot, R. (2018). Mismatch Unemployment and the Geography of Job Search. *American Economic Journal: Macroeconomics*, *10*(3), 42–70.

McDonald, J. F. & Moffitt, R. A. (1980). The Uses of Tobit Analysis. *The Review of Economics and Statistics*, *62*(2), 318–321.

Mora, R. & Reggio, I. (2015). Didq: A Command for Treatment-effect Estimation under Alternative Assumptions. *The Stata Journal: Promoting communications on statistics and Stata*, *15*(3), 796–808.

Mortensen, D. T. & Pissarides, C. A. (1994). Job Creation and Job Destruction in the Theory of Unemployment. *The Review of Economic Studies*, *61*(3), 397–415.

Pellizzari, M. (2010). Do Friends and Relatives Really Help in Getting a Good Job? *ILR Review*, *63*(3), 494–510.

Petrongolo, B. & Pissarides, C. A. (2001). Looking into the Black Box: A Survey of the Matching Function. *Journal of Economic Literature*, *39*(2), 390–431.

Pissarides, C. A. (1990). Equilibrium unemployment theory. Oxford, UK: Basil Blackwell.

Robins, J., Sued, M., Lei-Gomez, Q., & Rotnitzky, A. (2007). Comment: Performance of Double-Robust Estimators When "Inverse Probability" Weights Are Highly Variable. *Statistical Science*, *22*(4), 544–559.

Sant'Anna, P. H. C. & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, *219*(1), 101–122.

Smith, J. C. (2015). Pay Growth, Fairness, and Job Satisfaction: Implications for Nominal and Real Wage Rigidity. *The Scandinavian Journal of Economics*, *117*(3), 852–877.

Stops, M. (2016). Revisiting German labour market reform effects—a panel data analysis for occupational labour markets. *IZA Journal of European Labor Studies*, *5*(1).

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science : a review journal of the Institute of Mathematical Statistics*, *25*(1), 1–21.

Uusitalo, R. & Verho, J. (2010). The effect of unemployment benefits on re-employment rates: Evidence from the Finnish unemployment insurance reform. *Labour Economics*, *17*(4), 643–654.

Visser, J. (2006). Union membership statistics in 24 countries. *Monthly Labor Review*, *129*(1), 38–49.

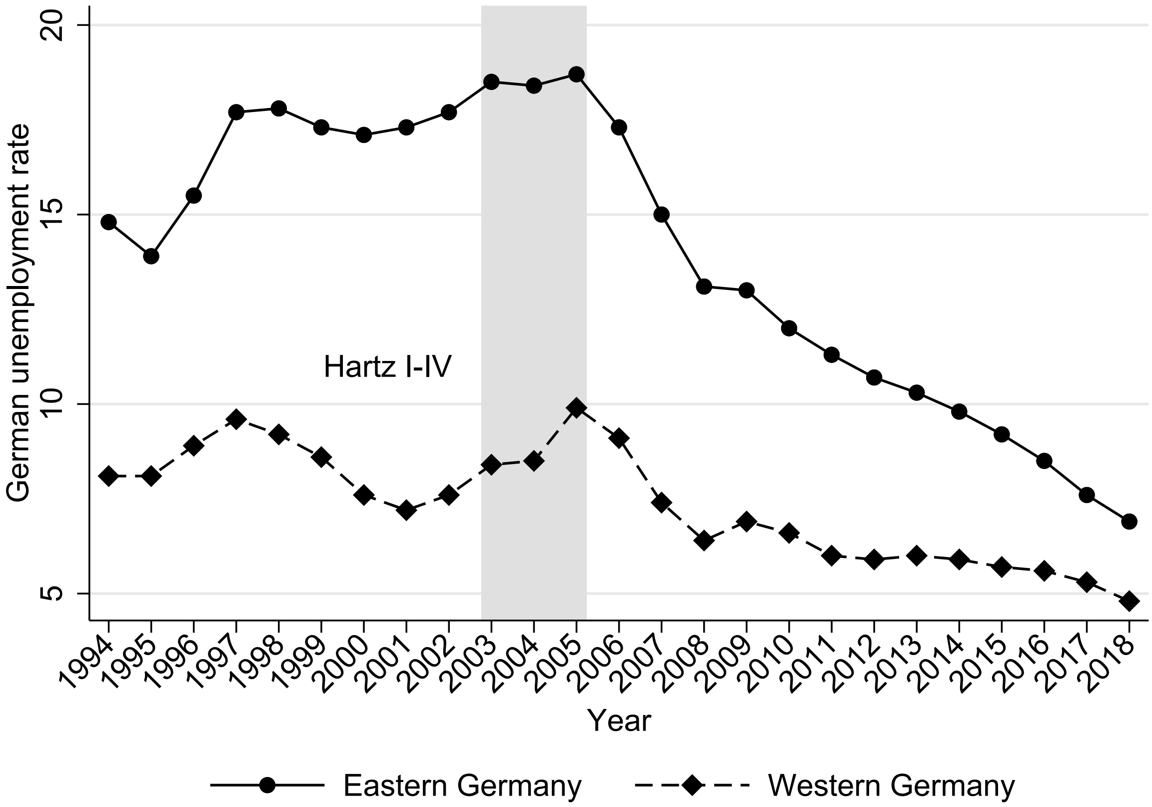
Wiese, R. (2014). What triggers reforms in OECD countries? Improved reform measurement and evidence from the healthcare sector. *European Journal of Political Economy*, *34*, 332–352.

Wooden, M. & Hawke, A. (2000). Unions and the Employment Growth: Panel Data Evidence. *Industrial Relations*, *39*(1), 88–107.

A Appendix

A.1 Figures

Figure A.1: German unemployment rate in Eastern and Western Germany



Notes: Data on the German unemployment rate is provided by the Federal Employment Agency time series: "Unemployment over time". The unemployment rate has its peak at roughly 11.1 percent in the year 2005 after which the rate falls remarkably. Trends are divided for Western and Eastern Germany. The grey shaded area marks the step-by-step introduction of the Hartz I–IV Reforms from 2003 to 2005.

A.2 Tables

Table A.1: Description and Explanation of Variables, $N = 14,658$

Variable	Description	Mean (Std. Dev.)
Dependent Variables		
Employment growth	Number of employees E_{it} in establishment i at year t and year $t-1$, divided by the average of employees in both years. $$g_{it} = \frac{E_{it} - E_{it-1}}{(E_{it} + E_{it-1})/2}$$	.025(.223)
Share of hires	Number of hires h_{it} in the year $t+1$ divided by the number of employees E_{it} in establishment i at year t. $$sh_{it} = \frac{100 * h_{it+1}}{E_{it}}$$	10.351(26.248)
Control Variables		
log(Employees)	Natural logarithm of the number of employees.	4.251(1.773)
log(Employees squared)	Natural logarithm of the squared number of employees.	21.212(15.506)
Pos. empl. expec.	Dummy variable equals 1 if establishment expects a positive employment trend in the next two years and 0 otherwise.	.309(.462)
Single establishment	Dummy variable equals 1 if the establishment is not part of a larger company or organization (i.e. single establishment) and 0 otherwise.	.619(.486)
Limited liability	Dummy variable equals 1 if the establishment is the legal form of a limited liability (e.g. GmbH, UG Ltd.) and 0 otherwise.	.613(.487)
Western Germany	Dummy variable equals 1 if the establishment is based in Western Germany and zero otherwise.	.741(.438)

Variable	Description	
Diverse ownership	Dummy variable equals 1 if the establishment has no dominant shareholder and 0 otherwise.	.059(.236)
Collective bargaining	Dummy variable equals 1 if the establishment is bound by an industry-wide wage agreement and 0 otherwise.	.501(.500)
Founded after 2000	Dummy variable equals 1 if the establishment was founded after the year 200 and 0 otherwise.	.239(.427)
Workforce Controls		
Share of female	Continuous measure for the share of female workers in relation to employment in year t	.412(.283)
Share of part-time	Continuous measure for the share of part-time workers in relation to employment in year t which require a university degree.	.192(.232)
Share of fixed-term	Continuous measure for the share of fixed-term workers in relation to employment in year t	.077(.158)
Share of high-skilled	Continuous measure for the share of high-skilled workers in relation to employment in year t which require a university degree.	.694(.272)
Share of apprentices	Continuous measure for the share of apprentices in relation to employment in year t which require a university degree.	.045(.082)

Table A.2: Results for OLS, FE and Tobit models

Dependent variable	Share of hires				Employment growth	
	OLS (1)	FE (2)	Tobit (3)	Het. Tobit (4)	OLS (5)	FE (6)
$HATRTZIII$	-3.45***	-6.11***	-1.76***	-.832***	-.020**	-.093***
	(.852)	(1.30)	(.450)	(.152)	(.008)	(.015)
$FEAuser$	1.74***		1.05***	.041	-.014***	
	(.586)		(.304)	(.105)	(.005)	
$HARTZIII \times FEAuser$	2.01**	2.03**	.927**	.417***	.025***	.037***
	(.791)	(.864)	(.414)	(.141)	(.007)	(.014)
(log) Employees	2.908***	28.701***	4.827***	2.731***	.078***	.723***
	(.486)	(1.160)	(.416)	(.238)	(.007)	(.067)
(log) Employees squared	-.360***	-2.469**	-.405***	-.214***	-.007***	-.046***
	(.050)	(1.042)	(.037)	(.021)	(.001)	(.007)
Single establishment	.677	-.405	.481**	.074	.016***	-.004
	(.473)	(1.242)	(.227)	(.076)	(.004)	(.013)
Limited liability	-.724	-1.102	-.305	.525***	-.015***	-.011
	(.556)	(1.405)	(.289)	(.119)	(.005)	(.017)
Share of part time employees	.209	4.623	.244	1.026**	.043***	-.058
	(1.470)	(3.435)	(.757)	(.422)	(.013)	(.036)
Share of female employees	-4.416***	5.191	-2.328***	-1.094***	-.001	.095*
	(1.256)	(6.105)	(.663)	(.358)	(.012)	(.052)
Share of qualified employees	-6.523***	2.215	-2.399***	-.893***	.009	-.014
	(1.235)	(4.657)	(.615)	(.251)	(.010)	(.032)
Share of fixed term employees	32.611***	21.849***	15.730***	14.181***	.043**	.190***
	(3.772)	(7.947)	(1.578)	(1.002)	(.021)	(.050)
Share of apprentices	-18.847***	-22.003***	-11.238***	-5.617***	.025	-.214*
	(1.948)	(7.538)	(1.312)	(.845)	(.028)	(.116)
Diverse ownership	-.804	1.851	-.086	.124	-.010	.008
	(.763)	(1.340)	(.393)	(.159)	(.008)	(.022)
Positive empl. expec.	3.215***	1.192	1.721***	1.032***	.050***	.029***
	(.527)	(.737)	(.267)	(.101)	(.004)	(.007)
Collective bargaining	-.033	.811	-.376	-.691***	-.018***	.016
	(.490)	(1.112)	(.246)	(.096)	(.004)	(.010)
Founded year $\geq$ 2000	7.778***	-1.182	4.538***	.642***	.067***	.026
	(.749)	(2.332)	(.411)	(.162)	(.006)	(.022)
Constant	9.59***	-73.26***	24.58***	16.04***	-.166***	-2.26***
	(2.64)	(25.44)	(1.46)	(.463)	(.028)	(.180)
Establishment fixed effects		✓				✓
Industry fixed effects	✓	✓	✓	✓	✓	✓
Federal state fixed effects	✓	✓	✓	✓	✓	✓
Year fixed effects	✓	✓	✓	✓	✓	✓
Control variables	✓	✓	✓	✓	✓	✓
R^2 / Pseudo R^2	.166	.861	.027	.034	.073	.820
Left (0) censored obs.			3,854	3,854		
Uncensored obs.			10,804	10,804		
Observations	14,658	14,658	14,658	14,658	14,658	14,658

Notes: IAB Establishment Panel, waves 2000–2008 with an overall sample size of $N = 14,658$ observations. Cluster-robust standard errors at the establishment level in parentheses. The control group consists of establishments that do not use the Federal Employment Agency, identified as reported in Section 4.2. The treatment group consists of establishments which do use the placement services. The latter group, therefore, is affected by the Hartz III reform which was implemented on January 1st, 2004. Estimation regarding the specification in Equation (3). Tobit model denotes the homoscedastic tobit model and in the heteroscedastic tobit model we include a vector of establishment size and industry dummy variables for the variance estimation. For more information regarding the heteroscedastic tobit model see Section 4.3. Year fixed effects include year dummy variables ranging from the years 2001 to 2008 with the year 2000 being the base category. Control variables are included as outlined in Section4.3. Fixed effects are nested within establishment cluster. *, **, *** significant at the 10%, 5%, 1% level.

Table A.3: Probit estimates for obtaining the propensity score to calculate inverse probability weights

Dependent Variable: FEA user	2000 (1)	2001 (2)	2002 (3)	2003 (4)	2004 (5)	2005 (6)	2006 (7)	2007 (8)	2008 (9)
(log) Employees	-.036	-.052	.012	.159	.121	.324***	-.042	.189**	.142*
	(.085)	(.084)	(.092)	(.098)	(.102)	(.103)	(.091)	(.081)	(.077)
(log) Employees squared	.017*	.021**	.014	.002	-.001	-.021*	.018*	-.003	-.001
	(.009)	(.009)	(.010)	(.011)	(.011)	(.011)	(.010)	(.009)	(.009)
Single establishment	-.016	.016	.017	.023	.005	.121	.037	.146**	.148**
	(.073)	(.070)	(.079)	(.094)	(.090)	(.090)	(.077)	(.068)	(.067)
Limited liability	-.149*	-.004	-.007	-.071	-.012	.032	.126	.045	-.054
	(.077)	(.080)	(.089)	(.108)	(.111)	(.117)	(.097)	(.085)	(.080)
Share of part time employees	-.094	-.128	-.391**	-.092	-.596***	-.242	-.386**	-.310*	-.428***
	(.192)	(.181)	(.192)	(.216)	(.225)	(.230)	(.189)	(.168)	(.153)
Share of female employees	-.208	-.081	.005	-.058	-.049	-.086	.024	.090	.020
	(.180)	(.167)	(.183)	(.210)	(.207)	(.225)	(.195)	(.161)	(.154)
Share of qualified employees	-.038	-.026	-.273*	-.044	.024	.243	-.217	-.124	-.022
	(.135)	(.130)	(.146)	(.174)	(.179)	(.189)	(.162)	(.139)	(.128)
Share of fixed term empl.	1.249***	1.270***	1.051***	.961***	.735***	1.162***	1.070***	.580***	.802***
	(.331)	(.297)	(.287)	(.282)	(.283)	(.282)	(.246)	(.210)	(.193)
Share of apprentices	1.340***	1.320***	.874*	2.394***	.266	1.624**	.146	1.066**	1.067**
	(.465)	(.394)	(.451)	(.618)	(.554)	(.648)	(.472)	(.485)	(.416)
Diverse ownership	.058	-.051	.100	-.065	.268	.102	-.108	.034	-.183
	(.147)	(.140)	(.160)	(.173)	(.178)	(.183)	(.147)	(.135)	(.123)
Positive empl. expec.	.124*	.125*	.019	-.050	-.077	.024	.083	.025	.035
	(.072)	(.070)	(.081)	(.094)	(.093)	(.097)	(.076)	(.065)	(.062)
Collective bargaining	.038	-.015	-.144*	-.068	.031	-.001	-.049	.054	.005
	(.073)	(.069)	(.080)	(.091)	(.092)	(.091)	(.077)	(.068)	(.065)
Founded year $\geq$ 2000	-.036	.102	-.091	.104	.076	.041	-.022	-.069	.110
	(.110)	(.096)	(.091)	(.101)	(.100)	(.103)	(.082)	(.070)	(.067)
Constant	.024	-.390	-.208	-.033	-.850**	-.799	-.220	-.688*	-.894**
	(.451)	(.399)	(.340)	(.531)	(.432)	(.503)	(.413)	(.373)	(.377)
Industry fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Federal state fixed effects	✓	✓	✓	✓	✓	✓	✓	✓	✓
Pseudo R^2	.119	.128	.123	.140	.130	.161	.117	.120	.119
Log-Likelihood	-1062.43	-1177.99	-961.18	-688.58	-705.94	-660.68	-959.08	-1194.99	-1310.64
Observations	1773	1979	1621	1179	1198	1144	1574	1986	2163

Notes: IAB Establishment Panel, waves 2000–2008 and an overall sample size of $N = 14,617$ observations. Robust standard errors in parentheses. Control variables are included as outlined in Section 4.3. Predicted propensity scores from these regressions are used in each cross-section to balance observations and adjust for observable differences as suggested by Imbens & Wooldridge (2009). The control group then receives the weights which are calculated as $w_t^c = \frac{1}{(1-p_t)}$ and the treatment group receive weights which are calculated as $w_t^t = \frac{1}{p_t}$. Significance: *, **, *** significant at the 10%, 5%, 1% level.

Table A.4: Mean differences of matched covariates using ATE weights

Year	Mean		t-Test		Mean		t-Test		Mean		t-Test
2000	Control	Treated	p-value	**2001**	Control	Treated	p-value	**2002**	Control	Treated	p-value
(log) Employees	4.364	4.380	.863		4.269	4.259	.914		4.140	4.124	.891
(log) Employees squared	21.846	22.069	.792		21.109	21.049	.940		20.408	20.312	.924
Single establishment	.624	.611	.607		.634	.632	.938		.642	.640	.951
Limited liability	.583	.575	.744		.596	.588	.758		.549	.540	.772
Share of part time employees	.155	.158	.782		.176	.177	.889		.201	.204	.810
Share of female employees	.410	.410	.992		.412	.415	.852		.430	.426	.823
Share of qualified employees	.637	.638	.930		.675	.672	.852		.644	.662	.332
Share of fixed term empl.	.054	.055	.889		.053	.055	.755		.090	.076	.444
Share of apprentices	.051	.048	.618		.047	.047	.986		.063	.054	.510
Diverse ownership	.056	.055	.929		.055	.061	.692		.052	.055	.790
Positive empl. expec.	.321	.331	.698		.272	.288	.516		.239	.243	.889
Collective bargaining	.569	.572	.908		.546	.542	.892		.502	.518	.611
Founded year $\geq$ 2000	.114	.116	.929		.119	.117	.929		.190	.189	.951

Notes: Weighted means for FEA user and FEA non-user. The control group receives the weights which are calculated as $w_t^c = \frac{1}{(1-p_t)}$ and the treatment group receive weights which are calculated as $w_t^t = \frac{1}{p_t}$. Here, p_t is the propensity score for each cross-section calculated as the predicted probability of receiving the treatment stemming from Probit estimates in Table A.3. Each last column displays the p-value of a t-test for the difference in mean values. Significance: *, **, *** significant at the 10%, 5%, 1% level.

Table A.5: **Mean differences of matched covariates using ATE weights: *continued***

Year	Mean		t-Test		Mean		t-Test		Mean		t-Test
2003	Control	Treated	p-value	**2004**	Control	Treated	p-value	**2005**	Control	Treated	p-value
(log) Employees	4.029	4.059	.815		4.390	4.391	.999		4.358	4.344	.920
(log) Employees squared	19.510	19.770	0.813		22.591	22.529	.960		22.386	22.422	.976
Single establishment	.657	.648	.764		.626	.620	.840		.556	.560	.903
Limited liability	.581	.579	.930		.638	.652	.663		.632	.628	.899
Share of part time employees	.183	.192	.538		.185	.175	.555		.172	.174	.857
Share of female employees	.421	.428	.688		.413	.408	.791		.416	.414	.902
Share of qualified employees	.709	.704	.785		.690	.697	.717		.713	.718	.800
Share of fixed term empl.	.068	.073	.671		.094	.087	.655		.099	.097	.930
Share of apprentices	.048	.047	.920		.049	.045	.691		.043	.043	.995
Diverse ownership	.050	.056	.671		.064	.063	.943		.068	.071	.900
Positive empl. expec.	.265	.270	.874		.254	.256	.936		.279	.289	.761
Collective bargaining	.483	.467	.656		.495	.506	.756		.515	.518	.924
Founded year $\geq$ 2000	.220	.225	.871		.243	.232	.712		.255	.270	.623

Notes: Weighted means for FEA user and FEA non-user. The control group receives the weights which are calculated as $w_t^c = \frac{1}{(1-p_t)}$ and the treatment group receive weights which are calculated as $w_t^t = \frac{1}{p_t}$. Here, p_t is the propensity score for each cross-section calculated as the predicted probability of receiving the treatment stemming from Probit estimates in Table A.3. Each last column displays the p-value of a t-test for the difference in mean values. Significance: *, **, *** significant at the 10%, 5%, 1% level.

Table A.6: Mean differences of matched covariates using ATE weights: *continued*

Year	Mean		t-Test		Mean		t-Test		Mean		t-Test
2006	Control	Treated	p-value	**2007**	Control	Treated	p-value	**2008**	Control	Treated	p-value
(log) Employees	4.253	4.204	.676		4.158	4.173	.872		4.041	4.020	.824
(log) Employees squared	21.473	20.954	.633		20.442	20.504	.937		19.458	19.350	.889
Single establishment	.628	.619	.742		.610	.610	.986		.630	.632	.948
Limited liability	.640	.624	.604		.657	.655	.959		.618	.321	.913
Share of part time employees	.209	.215	.720		.189	.192	.765		.213	.213	.952
Share of female employees	.408	.413	.781		.390	.392	.888		.416	.409	.653
Share of qualified employees	.729	.732	.878		.728	.725	.843		.708	.701	.620
Share of fixed term empl.	.079	.083	.754		.089	.089	.998		.093	.091	.904
Share of apprentices	.044	.042	.656		.041	.040	.818		.048	.044	.549
Diverse ownership	.062	.067	.788		.059	.059	.992		.062	.062	.978
Positive empl. expec.	.312	.308	.873		.387	.394	.780		.332	.332	.984
Collective bargaining	.454	.456	.955		.440	.446	.801		.457	.463	.806
Founded year $\geq$ 2000	.288	.280	.755		.329	.326	.906		.358	.350	.749

Notes: Weighted means for FEA user and FEA non-user. The control group receives the weights which are calculated as $w_t^c = \frac{1}{(1-p_t)}$ and the treatment group receive weights which are calculated as $w_t^t = \frac{1}{p_t}$. Here, p_t is the propensity score for each cross-section calculated as the predicted probability of receiving the treatment stemming from Probit estimates in Table A.3. Each last column displays the p-value of a t-test for the difference in mean values. Significance: *, **, *** significant at the 10%, 5%, 1% level.

Table A.7: Mean differences of matched covariates using ATT weights

Year	Mean		t-Test		Mean		t-Test		Mean		t-Test
2000	Control	Treated	p-value	**2001**	Control	Treated	p-value	**2002**	Control	Treated	p-value
(log) Employees	4.507	4.592	.417		4.451	4.510	.546		4.389	4.417	.826
(log) Employees squared	23.153	24.036	.374		22.713	22.332	.509		22.566	22.806	.844
Single establishment	.631	.599	.253		.630	.614	.551		.621	.628	.829
Limited liability	.584	.572	.694		.606	.609	.892		.549	.535	.709
Share of part time employees	.152	.154	.833		.169	.178	.422		.196	.201	.757
Share of female employees	.396	.393	.828		.401	.411	.564		.438	.431	.705
Share of qualified employees	.621	.636	.413		.666	.661	.732		.624	.658	.112
Share of fixed term empl.	.064	.065	.963		.064	.064	.971		.118	.090	.300
Share of apprentices	.059	.052	.502		.052	.053	.907		.075	.057	.364
Diverse ownership	.055	.052	.868		.050	.050	.959		.057	.061	.842
Positive empl. expec.	.324	.332	.768		.262	.282	.393		.216	.228	.658
Collective bargaining	.584	.602	.531		.569	.567	.956		.493	.524	.390
Founded year $\geq$ 2000	.109	.095	.450		.115	.122	.697		.171	.175	.869

Notes: Weighted means for FEA user and FEA non-user. The control group receives the weights which are calculated as $w_t^c = \frac{p_t}{(1-p_t)}$ and the treatment group receive weights which are calculated as $w_t^t = 1$. Here, p_t is the propensity score for each cross-section calculated as the predicted probability of receiving the treatment stemming from Probit estimates in Table A.3. Each last column displays the p-value of a t-test for the difference in mean values. Significance: *, **, *** significant at the 10%, 5%, 1% level.

Table A.8: Mean differences of matched covariates using ATT weights: *continued*

Year	Mean		t-Test		Mean		t-Test		Mean		t-Test
2003	Control	Treated	p-value	**2004**	Control	Treated	p-value	**2005**	Control	Treated	p-value
(log) Employees	4.351	4.383	.820		4.640	4.717	.591		4.580	4.733	.220
(log) Employees squared	22.150	22.433	.825		24.715	25.201	.739		24.008	25.335	.263
Single establishment	.645	.624	.697		.605	.599	.868		.553	.554	.827
Limited Liability	.583	.568	.697		.628	.643	.702		.634	.640	.885
Share of part time employees	.176	.193	.279		.182	.177	.797		.165	.175	.487
Share of female employees	.419	.433	.499		.417	.414	.898		.414	.409	.818
Share of qualified employees	.717	.705	.513		.689	.707	.397		.714	.722	.716
Share of fixed term empl.	.080	.089	.600		.117	.140	.572		.128	.119	.704
Share of apprentices	.057	.054	.705		.053	.048	.670		.047	.045	.739
Diverse ownership	.044	.058	.290		.072	.070	.922		.069	.063	.783
Positive empl. expec.	.231	.238	.810		.230	.248	.533		.258	.276	.594
Collective bargaining	.517	.490	.484		.523	.549	.483		.548	.560	.743
Founded year $\geq$ 2000	.202	.214	.683		.248	.233	.655		.244	.252	.788

Notes: Weighted means for FEA user and FEA non-user. The control group receives the weights which are calculated as $w_t^c = \frac{p_t}{(1-p_t)}$ and the treatment group receive weights which are calculated as $w_t^t = 1$. Here, p_t is the propensity score for each cross-section calculated as the predicted probability of receiving the treatment stemming from Probit estimates in Table A.3. Each last column displays the p-value of a t-test for the difference in mean values. Significance: *, **, *** significant at the 10%, 5%, 1% level.

Table A.9: Mean differences of matched covariates using ATT weights: *continued*

Year	Mean		t-Test		Mean		t-Test		Mean		t-Test
2006	Control	Treated	p-value	**2007**	Control	Treated	p-value	**2008**	Control	Treated	p-value
(log) Employees	4.493	4.536	.744		4.434	4.471	.699		4.280	4.362	.400
(log) Employees squared	23.745	23.802	.967		22.602	22.839	.786		21.366	21.888	.558
Single establishment	.628	.611	.571		.603	.601	.930		.622	.627	.864
Limited liability	.653	.611	.800		.682	.686	.878		.619	.658	.169
Share of part time employees	.203	.203	.963		.173	.186	.239		.199	.206	.577
Share of female employees	.413	.404	.612		.385	.391	.692		.421	.412	.596
Share of qualified employees	.728	.724	.791		.728	.721	.662		.709	.706	.838
Share of fixed term empl.	.098	.109	.445		.103	.103	.991		.112	.113	.942
Share of apprentices	.045	.044	.713		.046	.044	.788		.054	.047	.504
Diverse ownership	.054	.054	.999		.055	.056	.985		.052	.054	.837
Positive empl. expec.	.319	.311	.788		.376	.394	.492		.321	.342	.428
Collective bargaining	.460	.479	.535		.449	.468	.488		.473	.477	.864
Founded year $\geq$ 2000	.275	.280	.860		.304	.319	.546		.370	.373	.909

Notes: Weighted means for FEA user and FEA non-user. The control group receives the weights which are calculated as $w_t^c = \frac{p_t}{(1-p_t)}$ and the treatment group receive weights which are calculated as $w_t^t = 1$. Here, p_t is the propensity score for each cross-section calculated as the predicted probability of receiving the treatment stemming from Probit estimates in Table A.3. Each last column displays the p-value of a t-test for the difference in mean values. Significance: *, **, *** significant at the 10%, 5%, 1% level.

Table A.10: Distribution of establishments by German federal states

Federal State	Observations	Share
Schleswig-Holstein	651	4.441
Hamburg	659	4.496
Lower Saxony	1,122	7.655
Bremen	1,136	7.750
Nord Rhine-Westphalia	1,703	11.618
Hesse	1,149	7.839
Baden-Wuerttemberg	1,473	10.049
Bavaria	1,184	8.078
Saarland	642	4.380
Berlin	808	5.512
Brandenburg	618	4.216
Mecklenburg-West Pomerania	501	3.418
Saxony	838	5.717
Saxony-Anhalt	622	4.243
Thuringia	764	5.212
Rhineland-Palatinate	788	5.376
Total	14,658	100

Notes: IAB Establishment Panel, waves 2000 to 2008.

Table A.11: Distribution of establishments by size categories

Size category	Observations	Share
1-19	3,645	24.867
20-49	2,610	17.806
50-199	3,883	26.491
200-499	2,467	16.830
500+	2,053	14.006
Total	14,658	100

Notes: IAB Establishment Panel, waves 2000 to 2008.

Table A.12: **Distribution of establishments by IAB defined industries**

Industry classification (IAB Establishment Panel)	Observations	Share
Agriculture/forestry	169	1.153
Mining/energy	189	1.289
Food/luxury	381	2.599
Textiles/clothing	108	.737
Paper/printing	216	1.474
Wood sector	125	.853
Chemical sector	270	1.842
Plastics industry	256	1.746
Glass/stone/ore extraction	140	.955
Metal production	313	2.135
Recycling	20	0.136
Metal goods/steel production	551	3.759
Engineering	739	5.042
Vehicle engineering	208	1.419
Other vehicle production	91	0.621
Electrical engineering	359	2.449
Precision engineering/optics	236	1.610
Furniture/jewelry/toys	108	0.737
Main building sector	323	2.204
Building/installation	461	3.145
Car-rent/repairs/gas-stations	310	2.115
Wholesale trade	564	3.848
Retailing/repairs	618	4.216
Traffic	606	4.134
Telecommunications	47	0.321
Financial sector	336	2.292
Insurance	230	1.569
Data processing	305	2.081
Research/development	271	1.849
Judiciary/advertising	397	2.708
Realty/homes	124	0.846
Renting	1,286	8.773
Restaurants	500	3.411
Educational institutions	535	3.650
Health/social	1,567	10.690
Waste-management	68	0.464
Culture/sports/entertaining	169	1.153
Other services	260	1.774
Organizations	270	1.842
Civil service/social insurance	932	6.358
Total	14,658	100

Notes: IAB Establishment Panel, waves 2000 to 2008.