

Vouchers, Selection, and the Reform: Public Training Programs Revisited*

ULF RINNE[†]

IZA
Free University of Berlin

ARNE UHLENDORFF[‡]

IZA
DIW DC

ZHONG ZHAO[§]

IZA

April 28, 2008

Abstract

This paper studies the role of training vouchers in public training programs. Using a rich administrative data set from the Federal Employment Agency in Germany, we apply regression and propensity score matching procedures to measure the effect of the Hartz reform in 2003, which introduced training vouchers and imposed more selective criteria on the applicants. Besides estimating the total reform effect, we isolate the effect induced by changes in the composition of program participants (selection effect) from the effect based on the introduction of vouchers (voucher effect). Our results indicate that the selection effect plays only a minor role in explaining the generally positive impacts of the reform.

Keywords: Active Labor Market Policy; Program Evaluation; Propensity Score Matching; Voucher; Hartz Reform

JEL Classification: J64, J68, H43

*The data used in this paper originate from the evaluation of public training programs as part of the evaluation of the proposals of the Hartz Commission. Schneider et al. (2007) contains details. We would like to thank Hilmar Schneider, Marc Schneider, Marton Csillag, participants at the IZA Conference on the Evaluation of Labor Market Programs, the EALE Conference in Oslo, the EEA Meeting in Budapest, the CAPE in Nuremberg, the Conference of the German Statistical Society in Kiel and the VfS Conference in Munich for valuable discussions and helpful comments. All remaining errors are our own.

[†]Ulf Rinne, IZA, P.O. Box 7240, D-53072 Bonn. Email: rinne@iza.org

[‡]Arne Uhlendorff, IZA, P.O. Box 7240, D-53072 Bonn. Email: uhlendorff@iza.org

[§]Zhong Zhao, IZA, P.O. Box 7240, D-53072 Bonn. Email: zhao@iza.org

1 Introduction

Germany reformed its active labor market policy (ALMP) in 2003—known as the Hartz reform. One feature of the reform was the introduction of vouchers for public training programs. The former contracting-out system was abandoned and replaced by an innovative system. Job seekers are free to select their training provider in the market under the new regime, whereas this choice was previously made by the caseworker. Competition among providers should increase, and thereby program quality. While vouchers are a common instrument in other fields of public services, this approach is novel in the context of delivering ALMP.

From a theoretical point of view increased consumer choice and provider competition are the main arguments in favor of the new system. Consumers of training programs should be able to express their preferences optimally and thus maximize their utility. But there may be obstacles in case of public services which prevent from reaching higher levels of efficiency. For example, information asymmetries can lead to choices which do not truly reflect consumers' preferences. On the other hand, increased competition among providers is supposed to arise because of the opportunity for potential competitors to enter the market at little or no costs—the market then resembles a contestable market (Baumol, 1982). This could be the case in the post-reform regime, although for example reputation and brand names are potential barriers to entry for training providers which may remain.

Vouchers are a common instrument in other fields of public services—in particular in the field of education—and are quite extensively studied in the literature.¹ There are also studies on vouchers for pre-school education (e.g., Viitanen, 2007), but most of the examples focus on school education. For instance, Manski (1992) provides a theoretical model and presents simulation results on school vouchers and social mobility. Using a general equilibrium model, Nechyba (2000) studies private-school vouchers and residential mobility. Angrist et al. (2002) analyze the academic outcomes of students who win voucher lotteries. Ladd (2002) presents a review of major studies on school vouchers. More recently, Krueger and Zhu (2004) reexamine the New York City school voucher experiment, and Brunello and Checchi

¹The discussion about vouchers in the educational context started with Friedman (1962).

(2004) analyze the impact of vouchers for private schools introduced in some Italian regions. So far the overall picture that can be drawn from these studies is rather inconclusive, and the results are not very robust. For instance, studies with U.S. data typically indicate to have insufficient data to draw clear conclusions about the net effects on, e.g., student achievement or social and racial segregation. What can be learned—e.g., from large-scale programs in Chile or New Zealand—seems to be that large-scale universal school voucher programs do not generate substantial gains and could even be detrimental to sub-populations. On the other hand, more narrowly targeted programs seem to be more promising, but should be carefully implemented and only serve as one element of a broader strategy.

While school vouchers are quite extensively studied in the literature, there are only few studies of vouchers for adult education—mainly because vouchers are not very common in this context. Exceptions are mentioned in Dohmen (2003, 2007) which include the U.S. experience with the GI Bill and the Walloon training voucher scheme in Belgium. Another example is a World Bank initiative launched in Kenya in which training vouchers are given to entrepreneurs in the informal sector (see, e.g., Oosterbeek, 1998). The major difference between school vouchers and vouchers for adult education is that in the former case parents make the decision for their children, while in the latter situation participants make their own decision on how to use the voucher. The results on vouchers for adult education are in general rather positive: The GI Bill increased university enrollment (Levin, 2003), preliminary evidence suggests a successful start of the Belgian program (Geers, 2001), and vouchers had a beneficial impact on participants in Kenya although the impact on the training market was only modest (Hallberg, 2006).

A recent example for vouchers in the context of ALMP—although not in the field of education—is the job placement voucher. It was introduced in Germany in 2002 in order to end the public placement monopoly and subsidize private competitors. An increase in consumer sovereignty was also put forward as an argument in support of the voucher system. Winterhager et al. (2006) evaluate the effectiveness of this instrument and find a positive impact on the employment probability of voucher recipients in Western Germany. However, these vouchers are of little practical relevance as the take-up rate amounts to only about 5 percent receiving a

voucher. Bruttel (2005) gives some reasons for the little relevance (e.g., the complicated design, the low incentives for private providers).

This paper focuses on a voucher for adult education in the context of ALMP. We estimate the impact of the introduction of training vouchers on the effectiveness of public training programs. But next to this feature, other elements of the reform may also affect program effectiveness. These elements include a stricter selection rule for participants and a matching process between program types and participants by the caseworkers which is based on the expected reemployment probability. The overall effect of the Hartz reform could thus result from the introduction of the voucher system and/or from a change in the composition of participants because of the new selection criteria. We refer to the former as the voucher effect and to the latter as the selection effect. To decompose the overall reform effect into these two effects, we apply regression analysis and a two-step propensity score matching procedure to a rich administrative data set.

The remainder of this paper is organized as follows. Section 2 describes the institutional background of public training programs in Germany with a particular focus on changes between the pre- and post-reform period. After outlining the analytical framework in Section 3, we describe the data and the program types being analyzed in Section 4. Section 5 presents our results and, finally, Section 6 concludes.

2 Institutional Background

Germany's ALMP aims to increase the employment prospects of unemployed individuals. For this purpose, the Federal Employment Agency (FEA) spends a substantial amount of money on measures such as job creation schemes, public training programs, or employment subsidies. For instance, about 20.5 billion Euros were spent in 2002 (Eichhorst and Zimmermann, 2007). The most important part of ALMP in Germany are public training programs. With almost 7 billion Euros, these programs account for more than 32 percent of the expenditures. However, the number of participants decreased over the last years (see Figure 1). While more than 500,000 unemployed individuals entered a training program in 2000, this num-

ber approached only around 130,000 individuals in 2005. In 2006, it increased again to nearly 250,000 persons entering such programs.

[Figure 1 about here]

There already exists a number of studies evaluating the effectiveness of public training programs in Germany before the Hartz reform. For a recent review of the results see, e.g., Caliendo and Steiner (2005).² The results are quite heterogeneous—depending on the investigation period and the underlying data set. While earlier studies often find insignificant or even negative effects (Lechner, 1999, 2000; Hujer and Wellner, 2000), most of the recent studies which are based on rich administrative data sets find at least for some sub-groups positive treatment effects (Lechner et al., 2005, 2007; Fitzenberger et al., 2006; Biewen et al., 2007; Rinne et al., 2007). But there are also recent studies finding insignificant or negative effects (Hujer et al., 2006; Lechner and Wunsch, 2008). The mixed results may also be due to different methodological approaches. For instance, Stephan (2008) finds that estimated treatment effects differ considerably across different definitions of non-participation. The major lesson of the evaluation studies conducted for the pre-reform period seems to be that positive effects mainly occur in the longer run, and that studies which find positive medium- or long-term effects are also reporting negative short-term effects.

Before the Hartz reform, the provision of public training programs in Germany was organized as follows. After consultation with the job seeker, the caseworker in the local office of the FEA decided whether or not the unemployed individual should receive training. Courses were operated by private providers which were approved beforehand. The system is considered as a *de facto* contracting-out, although there were no legal contracts between providers and local FEA offices. Legally, job seekers paid the courses and were reimbursed, but usually the local offices paid the course fees directly to the providers in order to facilitate administration. The degree of competition among providers was only limited since approvals were granted only to exactly the number of providers needed to meet regional demand. A public tendering procedure was not in place. This informal procedure entailed a potential for collusive behavior between local FEA offices and private providers. For instance,

²The international literature on the evaluation of ALMP is summarized by Martin and Grubb (2001) and Kluge (2006), among others.

there was an informal guarantee that the capacity approved by the local office would be fully used. It was often reported that approved courses were simply filled up, even though the training provided was inappropriate for some individuals.

The provision of public training programs was substantially changed when the Hartz reform came into force on January 1, 2003. Next to changes which affected benefit claims during and after participation, the introduction of the training voucher (*Bildungsgutschein*) abandoned the former *de facto* contracting-out system.³ A training voucher is granted if the caseworker considers participation in a given type of public training program as a successful strategy to reintegrate the job seeker in the primary labor market—without taking into account the *relative* gain compared to the situation without participation. The selection criteria for participants thus became stricter after the reform; and the matching between program types and participants by the caseworkers which is also based on the expected reemployment probability is completely novel. The voucher—once it is granted—prescribes the program’s maximum duration, its intended educational target, its geographical scope, and the maximum course fee which will be reimbursed by the local FEA office. It is valid for three months. Within this period, job seekers are completely free to choose among approved training providers and courses in the market—subject to the requirements stated in the voucher.⁴ Local FEA staff are not allowed to make recommendations, but can provide, e.g., a list of approved courses. There was, however, a transitional arrangement when the reform was introduced: The allocation of participants into public training programs was *exclusively* based on vouchers only from March 2003 onwards (Schneider et al., 2007).⁵

Although the innovative voucher system should both increase consumer sovereignty and competition among training providers, Bruttel (2005) presents initial evidence that there are practical obstacles to fully achieve this positive effects. For instance, information asymmetries constrain consumer sovereignty. In particular low-qualified job seekers lack the abilities to navigate the training market and to take an active

³For details on changes concerning benefit claims see, e.g., Schneider et al. (2007).

⁴The approval of providers and courses is subject to a new quality management system which adopts a two-level approach. For details see, e.g., Bruttel (2005).

⁵The official transitional arrangement was as follows: “Individuals who were counselled before January 1, 2003 and participation in a public training program was agreed upon do not receive a training voucher if they enter the program until February 28, 2003.”

role in searching for an appropriate course. A potential obstacle for competition among providers is their unequal distribution across German regions. Additionally, it can be observed that providers react to the reform by increased co-operation or collusive behavior, e.g., they do not offer the same courses anymore.

This initial evidence is supplemented by Schneider et al. (2007) who analyze the implementation of the reform as a whole. Accordingly, the impacts of the reform primarily materialize in two dimensions. First, the composition of participants is affected. Participants in the post-reform period exhibit on average better employment prospects than in the pre-reform period. Second, the structure of program types is affected. The focus shifts towards regions with lower unemployment rates, courses with comparatively shorter durations, and courses providing qualifications and skills which fit regional short-term market demand. Moreover, Schneider et al. (2007) present some descriptive evidence on the receipt of the education voucher. While only about 30 percent of the participants who entered a program in the first quarter of 2003 received a voucher, the take-up rate increases to 70–80 percent in subsequent quarters. This is in line with the above mentioned transitional arrangement.

Considering the practical obstacles and the actual implementation, the reform impact on the effectiveness of public training programs is anything but clear. However, Schneider and Uhlenborff (2006) find that the effectiveness increases after the reform. Nonetheless, the question which features of the reform cause this increase—and to what particular extent—remains unanswered. We answer this question by decomposing the reform into two effects: (a) selection effect, and (b) voucher effect.

The selection effect stems from a different composition of participants between the pre- and the post-reform period. It is due to stricter selection criteria as well as to the *unintended* effect of the voucher that low-qualified job seekers lack the abilities to navigate the training market and to take an active role in searching for an appropriate course.

On the other hand, the voucher effect comprises the *intended* effects of the introduction of training vouchers. These effects include a potentially better match between participants and courses, a more market-oriented, i.e. demand-oriented, approach of the local FEA offices, and quality enhancements based on increased

competition among providers.

3 Analytical Framework

The aim of this paper is to estimate the effect of vouchers for publicly financed training programs, and to test the hypothesis that their introduction improves the quality (i.e., the effectiveness) of the programs under consideration. However, it is difficult to isolate the impact of vouchers since along with the introduction of the voucher system, there are other major changes going on as described above—especially a stricter selection rule for participants. The reform effect could thus result from a change of program quality and from a change of the composition of participants.⁶

A simple model to capture these two aspects and to isolate the voucher effect is the following:

$$Y_i = \alpha + X_i\beta + \delta D_i + \gamma R_i + \eta(D_i \times R_i) + \epsilon_i \quad (1)$$

where D_i is a dummy variable that takes the value one if individual i participates in the program and zero otherwise. R_i indicates the pre- and post-reform period similarly. Assuming that (i) there is no selection on unobservables, (ii) the treatment effect is homogeneous, and (iii) the outcome and covariates have a linear relationship, the coefficient η in equation (1) represents the impact of the voucher. In this case, controlling for X_i is equivalent to controlling for the compositions of training participants before and after the reform.⁷

However, in order to avoid parametric assumptions and to accommodate heterogeneity of treatment effects, we also apply propensity score matching procedures to isolate the impact of vouchers. Mueser et al. (2007) present evidence that if administrative data is used to measure the performance of training programs, propensity

⁶Changes in the general economic situation may be another component of the reform effect. However, in what follows we control for this issue as participants and matched non-participants are subject to the same cyclical environment.

⁷We run this regression on the raw sample of participants and non-participants separately for each program type twice: (i) *without* controlling for individual characteristics X_i ; and (ii) controlling for individual characteristics X_i . The results of these regression are discussed in Section 5.1 below.

score matching is generally most effective.

Using the potential outcome framework as in Rubin (1974), we assume that each individual has two potential outcomes for the program: Y_{1i} is the outcome if individual i participates, and Y_{0i} if not. Let D_i be again an indicator for participation, we can define different treatment effects in a similar way as Heckman and Vytlacil (1999, 2005):

$$TE_i = Y_{1i} - Y_{0i} \quad (\text{Treatment effect for individual } i)$$

$$ATE = E[TE_i] \quad (\text{Average treatment effect for the population})$$

$$ATT = E[TE_i | D_i = 1] \quad (\text{Average treatment effect on the treated})$$

and the average treatment effects on the treated before and after the reform are

$$ATT_b = E[TE_i | D_i = 1, R_i = 0] \quad (\text{ATT pre-reform period})$$

$$ATT_a = E[TE_i | D_i = 1, R_i = 1] \quad (\text{ATT post-reform period})$$

Under the matching assumptions of Rosenbaum and Rubin (1983)

$$(Y_{0i}, Y_{1i}) \perp D_i | X_i \quad (\text{Conditional independence assumption})$$

$$0 < \text{prob}(D_i = 1 | X_i) < 1 \quad (\text{Common support assumption})$$

ATT_b (or ATT_a) can be estimated from pre-reform data (or post-reform data) by propensity score matching methods and it corresponds to δ (or $\delta + \eta$) in equation (1).⁸

However, the difference between ATT_b and ATT_a does *not* equal the effect of the introduction of vouchers, since the participants before and after reform may have different characteristics. As mentioned above, compared to the pre-reform period, the post-reform programs are more selective (possibly leading to a selection effect, SE) and vouchers are introduced (which may cause a voucher effect, VE).

Hence,

$$ATT_a = ATT_b + VE + SE \quad (5)$$

⁸Heckman et al. (1997, 1998) present a weaker version of the conditional independence assumption: $E[Y_{0i} | D_i = 0, X_i] = E[Y_{0i} | D_i = 1, X_i]$.

and the total reform effect (RE) can be written as:

$$\begin{aligned} RE &= ATT_a - ATT_b \\ &= VE + SE \end{aligned} \tag{6}$$

To be able to isolate the voucher effect, we apply a two-step propensity score matching procedure. In the first step, pre-reform participants are matched with post-reform participants. As a result, the obtained pairs of participants only differ with respect to the timing of participation. Importantly, observable characteristics do not differ anymore. In the second step, the matched pre-reform participants in 2002 are matched with non-participants of the same year. The corresponding treatment effect is the effect *only* for those participants under the pre-reform regime who are comparable to participants after the reform (we henceforth refer to this effect as ATT_{bQ}), so this step controls for the selection effect.

With ATT_{bQ} we can calculate the difference in differences of the treatment effects to estimate the voucher effect:

$$VE = ATT_a - ATT_{bQ} \tag{7}$$

Finally, the comparison of the voucher effect with the reform effect gives us an estimate of the selection effect:

$$\begin{aligned} SE &= RE - VE \\ &= (ATT_a - ATT_b) - (ATT_a - ATT_{bQ}) \\ &= ATT_{bQ} - ATT_b \end{aligned} \tag{8}$$

4 Data

We use a sample of a particularly rich administrative data set, the Integrated Employment Biographies (IEB) of the FEA.⁹ It contains detailed daily information on

⁹The IEB is in general not publicly available. Only a 2.2 percent random sample (the Integrated Employment Biographies Sample, IEBS) can be obtained for research purposes. See, e.g., Hummel et al. (2005) for details on the IEBS. The IEB consists of four different administrative data sources: the employees' history (BeH), the benefit recipients' history (LeH), the job seekers' data base (ASU/BewA), and the program participants' master data set (MTH). For a detailed description

employment subject to social security contribution including occupational and sectoral information, receipt of transfer payments during periods of unemployment, job search, and participation in different programs of ALMP. Furthermore, the IEB comprises a large variety of covariates—e.g., age, marital status, number of dependent children, disability, nationality and education.

Since the public training programs currently in place in Germany are quite heterogeneous, we concentrate on and differentiate between four particular types: (a) type 1: occupation-related or general training, (b) type 2: practice training in key qualifications, (c) type 3: practice firm, and (d) type 4: group training with occupation-related certificate.

Participants in type 1 learn specific skills required for a certain vocation (e.g., computer-aided design for a technician/tracer) or receive qualifications that are of general vocational use (e.g., MS Office, computer skills). Type 2 is a predominantly practically oriented program with only few theoretical parts. It follows the principle ‘learning by doing’ and is often combined with internships. Within type 3 the simulation of real operations is conducted, and most of the times technical training is provided. For example, participants are endowed with practical skills of wood working and processing at work benches and machines under the supervision of instructors. Type 4 is a group training measure aiming to provide an occupation-related certificate. More specifically, a group of participants attends the same retraining measure at an educational institution. The measure also includes periods of practical training in certified companies/organizations. The aim is to provide participants a vocational degree by passing an examination.

Figure 2 shows that type 1 is by far the most important program type. In the pre-reform period, about 60 percent of all participants in public training programs were assigned to this particular type. It became even more important after the reform in 2003 as this share increased to more than 70 percent. Moreover, the four types together account for roughly 90–95 percent of all participants in public training programs over the period 2000–2004.

[Figure 2 about here]

see, e.g., Schneider et al. (2007).

As Figure 3 indicates, the program types 1–3 are—in comparison to other ALMP measures in Germany—rather shorter measures. Both in the pre-reform and in the post-reform period, after one year more than 90 percent of the participants have left each type. However, the program duration decreased after the reform was introduced, in particular as far as type 1 is concerned. While the median program duration is about 8 months for this program type in the pre-reform period, it amounts to about 6 months after the reform. The median duration for types 2 and 3 decreased only slightly and remained rather stable at about 6 months. In contrast, program type 4 constitutes a rather long-term program. Its median duration is 24 months in the pre-reform period. In case of this program type, we observe the largest decline in program duration.

[Figure 3 about here]

Our sample of participants consists of 275 unemployed persons per quarter and program for the year 2003, i.e., we observe around 1,100 participants for each program type after the introduction of the training vouchers. In order to apply the matching approach as described in section 2, 20 participants from the period before the reform were drawn per participant in 2003, i.e., we draw a sample of 22,000 participants in 2002 per program type.

Beyond the matching of participants after with participants before the reform, we need to match participants with non-participants. Therefore, we draw 80 non-participants per participants. These individuals had to be unemployed for the same duration as the corresponding participants. Moreover, they are required to not having participated in the respective type of training before and in the quarter of the participant's program entry, but we do not condition on future non-participation.¹⁰

The success of program participation is evaluated by looking at the probability of being employed starting at the (fictitious) program entry over a period of 18 months. This period is based on the fact that we focus on program participation in the years 2002 and 2003, and we can observe reliable data for all employment

¹⁰In contrast, e.g., Lechner and Wunsch (2008) require non-participation in the follow-up period after the (fictitious) program entry for comparison individuals. Although we opted for the above stated definition of non-participation and do not exclude future participants, the alternative approach clearly has the advantage of employing a very straightforward definition of non-participation.

states until June 30, 2005. Individuals are regarded as employed if they hold a job in the primary labor market. For instance, participation in job creation schemes is not included in our outcome measure. Moreover, the administrative data set only includes employment that is subject to social security contributions.¹¹

5 Results

In this section, we report the results based on the two approaches outlined above: (i) regression analysis, and (ii) two-step matching procedure. Estimates of the differences in employment probabilities are presented for a period of two years after the (fictitious) program entry.¹²

5.1 Regression Analysis

We run the regression according to equation (1) on the raw sample of participants and non-participants separately for each program type twice: (i) *without* controlling for individual characteristics X_i ; and (ii) controlling for individual characteristics X_i . We present estimates on the probability of being employed 6, 12, and 18 months after the program entry, respectively. The results of these regression are displayed in Table 1.

[Table 1 about here]

As mentioned above, our primary interest lies on the estimate for the coefficient η , which represents the impact of the voucher. For all program types being analyzed, we consistently find significantly positive voucher effects on the outcome variable 6 and 12 months after the program entry, respectively. Moreover, these results are quite robust to controlling for individual characteristics X_i .¹³ However, the voucher

¹¹This means that, e.g., we do not observe self-employment, and remunerations are only reported up to the social security contribution ceiling.

¹²We thus follow the prevailing approach in the recent evaluation literature. A different approach concentrates on treatment effects only after the end of the program. For advantages and disadvantages of both approaches see, e.g., Caliendo and Kopeinig (2008).

¹³Type 2 constitutes an exception in this context. After controlling for for individual characteristics X_i , the voucher effect becomes negative and/or insignificant.

effects turn out to be in general significantly negative 18 months after program entry, respectively. Only for program type 4, the voucher effect remains significantly positive.

5.2 Two-Step Matching Procedure

By applying the described two-step matching procedure, we are able to decompose the reform effect into two separate components: a voucher effect and a selection effect. For this purpose, the treatment effects for the respective matched samples are calculated as the difference in mean outcomes between the matched groups. Below, we present estimates of differences in employment probabilities for a period of 18 months after program entry, calculated every fortnight.¹⁴

Firstly, we display estimates of the reform effect in Figure 4. The underlying ATT_a and ATT_b are based on differences in mean outcomes between program entrants in 2002 for the pre-reform period (in 2003 for the post-reform period) and matched non-participants *without* controlling for potential changes in the composition of participants between the two periods. The reform effect is then simply the difference between the two depicted lines.

[Figure 4 about here]

We observe that participants in all analyzed program types face a substantial lock-in effect.¹⁵ In the first months after entering the program, the employment probabilities of participants are considerably lower than those of matched non-participants. The duration and the extent of these lock-in effects vary by program type. While the rather shorter program types 1–3 exhibit relatively moderate lock-in effects, the longer-term program type 4 leads to a considerable reduction in employment probabilities for the whole observation period. In contrast, the treatment effects for program types 1–3 become generally positive about 12 months after the program entry.

¹⁴The matching algorithms are implemented using the PSMATCH2 Stata ado-package by Leuven and Sianesi (2003).

¹⁵While participating—or being ‘locked-in’ in the program—individuals probably reduce their search activities for new jobs (van Ours, 2004).

The comparison of the estimates between the pre- and the post-reform period in Figure 4 reveals—if anything—positive reform effects for the program types being analyzed. However, the differences between ATT_a and ATT_b are very small for types 2 and 3. On the other hand, we observe significantly positive reform effects for types 1 and 4—at least for parts of the observation period. In particular, the lock-in effect of program type 4 seems to be considerably reduced after the reform.

In Figure 5 we present estimates of the voucher effect. The underlying ATT_a and ATT_{bQ} are based on differences in mean outcomes between *matched* program entrants in 2002 for the pre-reform period (program entrants in 2003 for the post-reform period) and matched non-participants, respectively. By doing so, we take potential changes in the composition of participants between the two periods into account. The voucher effect is then the difference between the two depicted lines.

[Figure 5 about here]

In short, Figure 5 looks quite similar to Figure 4. This is not surprising for the post-reform period, since we again depict ATT_a for this period. However, the differences between ATT_b and ATT_{bQ} seem to be rather small.

Figure 6 reveals more insights about the extent and magnitude of reform effects, voucher effects, and selection effects. It displays the decomposition of the reform effect, and thus summarizes the previous graphs.

[Figure 6 about here]

The decomposition shows that with respect to types 1, 2 and 3, the reform effect seems to be almost exclusively based on the voucher effect. For most of the observation period, we observe a negative selection effect in case of these three program types. On the other hand, the selection effect seems to be a more important driving force behind the reform effect as far as type 4 is concerned. In this case, the selection effect becomes positive about 6 months after program entry. 18 months after program entry, the selection effect roughly accounts for half of the reform effect.

The results of two-step matching procedures thus indicate that the selection effect only plays a minor role in explaining the generally positive impacts of the

reform in 2003—in particular, as far as the shorter program types are concerned. For these program types, the voucher effect seems to raise program quality, leading to an increased effectiveness. This finding is consistent with Lechner and Smith (2007). An exception to this general statement of negligible selection effects constitutes the longer-term program type 4, for which the selection effect indeed seems to be an important factor for the increased effectiveness in the post-reform period. On the other hand, the voucher effect still accounts for about 50 percent of the increased effectiveness we observe for this program type.

6 Conclusion

This paper analyzes the impacts of the labor market reform in 2003 on the effectiveness of publicly financed training in Germany. Using a rich administrative data set from the FEA, we apply regression and propensity score matching procedures to address the question whether the introduction of vouchers raised the quality of the programs by means of an increased competition. Therefore, we decompose the total reform effect into two separate components. By distinguishing between a voucher effect and a selection effect, we disentangle the effects that are based on the introduction of vouchers from those that are based on changes in the composition of program participants.

Regression analysis reveals generally positive voucher effects on the outcome variable 6 and 12 months after program entry. However, voucher effects turn out to be in general negative 18 months after program entry, respectively. Only for program type 4, the voucher effect remains positive.

Our results obtained from the two-step matching procedure indicate that the selection effect plays only a minor role in explaining the generally positive impacts of the Hartz reform. In particular for shorter training programs, the voucher effect seems to raise program quality, leading to an increased effectiveness. An exception to this general statement of negligible selection effects constitutes the longer-term program type 4, for which the selection effect indeed seems to be an important factor behind the increased effectiveness in the post-reform period. On the other hand, the voucher effect still accounts for about half of the increased effectiveness.

In sum, we present evidence for weak selection effects. Only the effectiveness of the longer-term program type 4 seems to be positively affected by the strategy to increasingly provide training to individuals with better employment prospects.

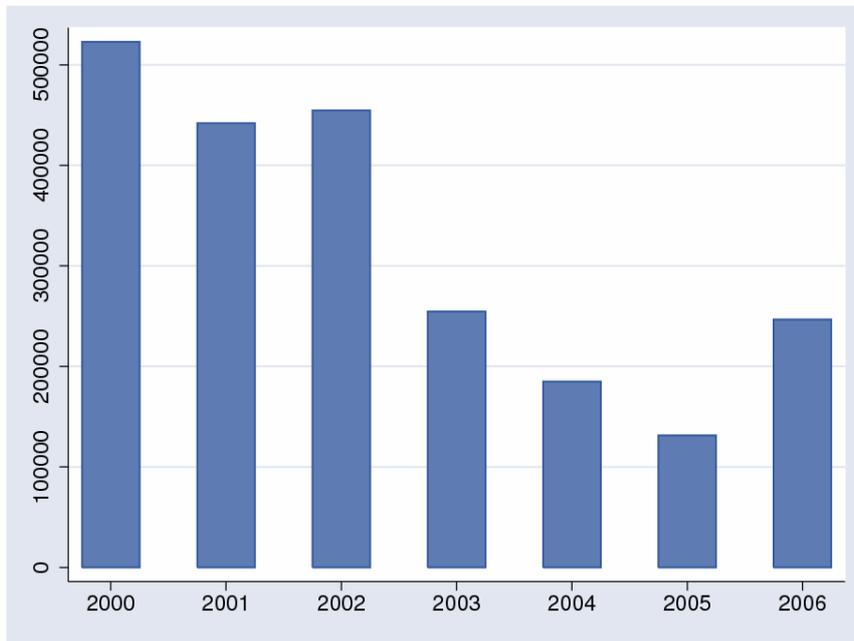
References

- Angrist, J., E. Bettinger, E. Bloom, E. King, and M. Kremer (2002). Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment. *American Economic Review* 92(5), 1535–1558.
- Baumol, W. (1982). Contestable Markets: An Uprising in the Theory of Industry Structure. *American Economic Review* 72(1), 1–15.
- Biewen, M., B. Fitzenberger, A. Osikominu, and M. Waller (2007). Which Program for Whom? Evidence on the Comparative Effectiveness of Public Sponsored Training Programs in Germany. *IZA Discussion Paper* 2885.
- Brunello, G. and D. Checchi (2004). School Vouchers Italian Style. *Giornale degli Economisti e Annali di Economia* 63(3/4), 357–399.
- Bruttel, O. (2005). Delivering Active Labor Market Policy through Vouchers: Experiences with Training Vouchers in Germany. *International Review of Administrative Sciences* 71(3), 391–404.
- Caliendo, M. and S. Kopeinig (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys* 22(1), 31–72.
- Caliendo, M. and V. Steiner (2005). Aktive Arbeitsmarktpolitik in Deutschland: Bestandsaufnahme und Bewertung der mikroökonomischen Evaluationsergebnisse. *Zeitschrift für ArbeitsmarktForschung* 38(2-3), 396–418.
- Dohmen, D. (2003). Alternative Formen der Finanzierung beruflicher Weiterbildung. *FiBS-Forum* 15.
- Dohmen, D. (2007). Current Trends in the of Demand-led Financing of Further Training in Europe: A Synopsis. *FiBS-Forum* 40.
- Eichhorst, W. and K. F. Zimmermann (2007). And Then There Were Four... How Many (and Which) Measures of Active Labor Market Policy Do We Still Need? *Applied Economics Quarterly* 53(3), 243–272.
- Fitzenberger, B., A. Osikominu, and R. Völter (2006). Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany. *IZA Discussion Paper* 2121.
- Friedman, M. (1962). The Role of Government in Education. In *Capitalism and Freedom*, Chapter 6. Chicago: University of Chicago Press.
- Geers, F. (2001). Belgium, Wallonia Region: The Training Voucher Scheme. In T. Pukkinen, C. Romijn, and S. Elson-Rogers (Eds.), *Funding Continuing Training in Small and Medium-Sized Enterprises. Discussion and Case Studies from Across the EU*, Volume 17 of *Cedefop Panorama Series*. Luxembourg.
- Hallberg, K. (2006). A Retrospective Assessment of the Kenya Voucher Training Program. *Small Enterprise Development* 17(2), 56–67.
- Heckman, J. and E. Vytlacil (1999). Local Instrumental Variables and Latent Variables Models for Identifying and Bounding Treatment Effects. *Proceedings of the National Academy of Sciences of the USA* 96, 4730–4734.

- Heckman, J. and E. Vytlacil (2005). Structural Equations, Treatment Effects and Econometric Policy Evaluation. *Econometrica* 73(3), 669–738.
- Heckman, J. J., H. Ichimura, and P. E. Todd (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *Review of Economic Studies* 64(4), 605–654.
- Heckman, J. J., H. Ichimura, and P. E. Todd (1998). Matching as an Econometric Evaluation Estimator. *Review of Economic Studies* 65(2), 261–294.
- Hujer, R., S. L. Thomsen, and C. Zeiss (2006). The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany. *Allgemeines Statistisches Archiv* 90(2), 299–322.
- Hujer, R. and M. Wellner (2000). Berufliche Weiterbildung und individuelle Arbeitslosigkeitdauer in West- und Ostdeutschland: Eine mikroökonomische Analyse. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 33(3), 405–419.
- Hummel, E., P. Jacobebbinghaus, A. Kohlmann, M. Oertel, C. Wübbecke, and M. Ziegerer (2005). Stichprobe der Integrierten Erwerbsbiographien, IEBS 1.0. *FDZ Datenreport* 6.
- Kluve, J. (2006). The Effectiveness of European Active Labor Market Policy. *IZA Discussion Paper* 2018.
- Krueger, A. B. and P. Zhu (2004). Another Look at the New York City School Voucher Experiment. *American Behavioral Scientist* 47(5), 658–698.
- Ladd, H. F. (2002). School Vouchers: A Critical Review. *Journal of Economic Perspectives* 16(4), 3–24.
- Lechner, M. (1999). Earnings and Employment Effects of Continuous Off-the-Job Training in East-Germany after Unification. *Journal of Business & Economic Statistics* 17(1), 74–90.
- Lechner, M. (2000). An Evaluation of Public-Sector-Sponsored Continuous Vocational Training Programs in East Germany. *The Journal of Human Resources* 35(2), 347–375.
- Lechner, M., R. Miquel, and C. Wunsch (2005). Long-Run Effects of Public Sector Sponsored Training in West Germany. *IAB Discussion Paper* 3.
- Lechner, M., R. Miquel, and C. Wunsch (2007). The Curse and Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany After Unification. *German Economic Review* 8(4), 468–509.
- Lechner, M. and J. A. Smith (2007). What is the Value Added by Caseworkers? *Labour Economics* 14(2), 135–151.
- Lechner, M. and C. Wunsch (2008). What Did All the Money Do? On the General Ineffectiveness of Recent West German Labour Market Programmes. *Kyklos* 61(1), 134–174.
- Leuven, E. and B. Sianesi (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. Statistical Software Components, Boston College Department of Economics. Available at <http://ideas.repec.org/c/boc/bocode/s432001.html>.

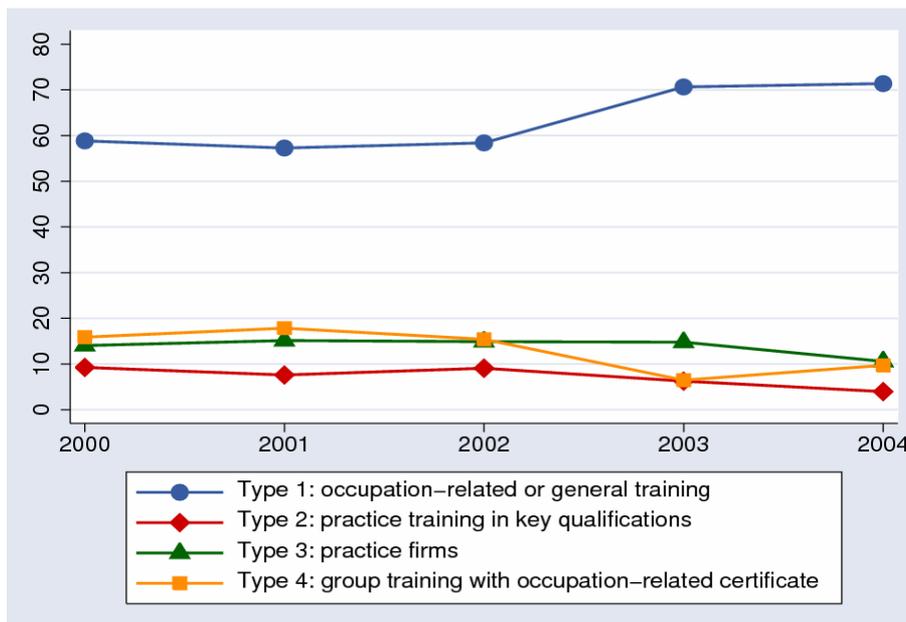
- Levin, H. M. (2003). Post-Compulsory Entitlements: Vouchers for Lifelong Learning. In D. Dohmen and B. A. Cleuvers (Eds.), *Finanzierung von Weiterbildung und lebenslangem Lernen*, Volume 2 of *Schriften zur Bildungs- und Sozialökonomie*. Bielefeld.
- Manski, C. F. (1992). Educational Choice (Vouchers) and Social Mobility. *Economics of Education Review* 11(4), 351–369.
- Martin, J. P. and D. Grubb (2001). What Works and for Whom: A Review of OECD Countries' Experiences with Active Labor Market Policies. *Swedish Economic Policy Review* 8(2), 9–56.
- Mueser, P. R., K. R. Troske, and A. Gorislavsky (2007). Using State Administrative Data to Measure Program Performance. *The Review of Economics and Statistics* 89(4), 761–783.
- Nechyba, T. J. (2000). Mobility, Targeting, and Private-school Vouchers. *American Economic Review* 90(1), 130–146.
- Oosterbeek, H. (1998). Innovative Ways to Finance Education and Their Relation to Lifelong Learning. *Education Economics* 6(3), 219–251.
- Rinne, U., M. Schneider, and A. Uhlendorff (2007). Too Bad to Benefit? Effect Heterogeneity of Public Training Programs. *IZA Discussion Paper* 3240.
- Rosenbaum, P. R. and D. B. Rubin (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70(1), 41–55.
- Rubin, D. B. (1974). Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies. *Journal of Educational Psychology* 66(5), 688–701.
- Schneider, H., K. Brenke, B. Jesske, L. Kaiser, U. Rinne, M. Schneider, J. Steinwede, and A. Uhlendorff (2007). Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission – Bericht 2006. *IZA Research Report* 10.
- Schneider, H. and A. Uhlendorff (2006). Die Wirkung der Hartz-Reform im Bereich der beruflichen Weiterbildung. *Zeitschrift für ArbeitsmarktForschung* 39(3-4), 477–490.
- Stephan, G. (2008). The Effects of Active Labor Market Programs in Germany: An Investigation Using Different Definitions of Non-treatment. *IAB Discussion Paper* 12.
- van Ours, J. C. (2004). The Locking-in Effect of Subsidized Jobs. *Journal of Comparative Economics* 32(1), 37–55.
- Viitanen, T. (2007). Public versus Private Provision of Daycare: An Experimental Evaluation. *IZA Discussion Paper* 3009.
- Winterhager, H., A. Heinze, and A. Spermann (2006). Deregulating Job Placement in Europe: A Microeconomic Evaluation of an Innovative Voucher Scheme in Germany. *Labour Economics* 13(4), 505–517.

Figure 1: Number of Entrants into Public Training Programs (2000–2006).



Source: Federal Employment Agency (FEA).

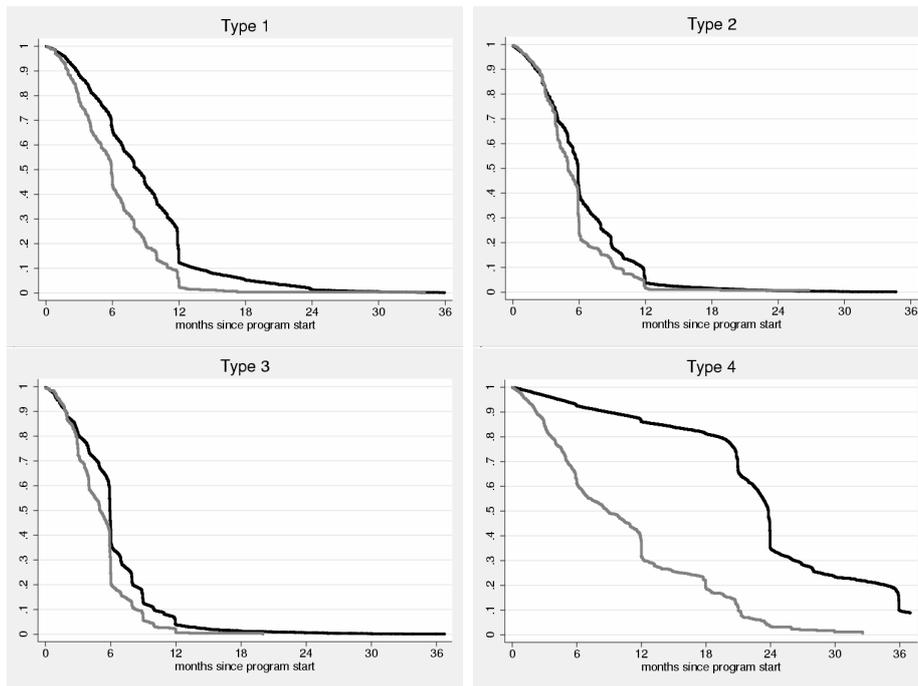
Figure 2: Share of Entrants by Program Type (2000–2004).



Source: Federal Employment Agency (FEA).

Note: Share (in percent) in annual number of entrants in publicly financed training programs.

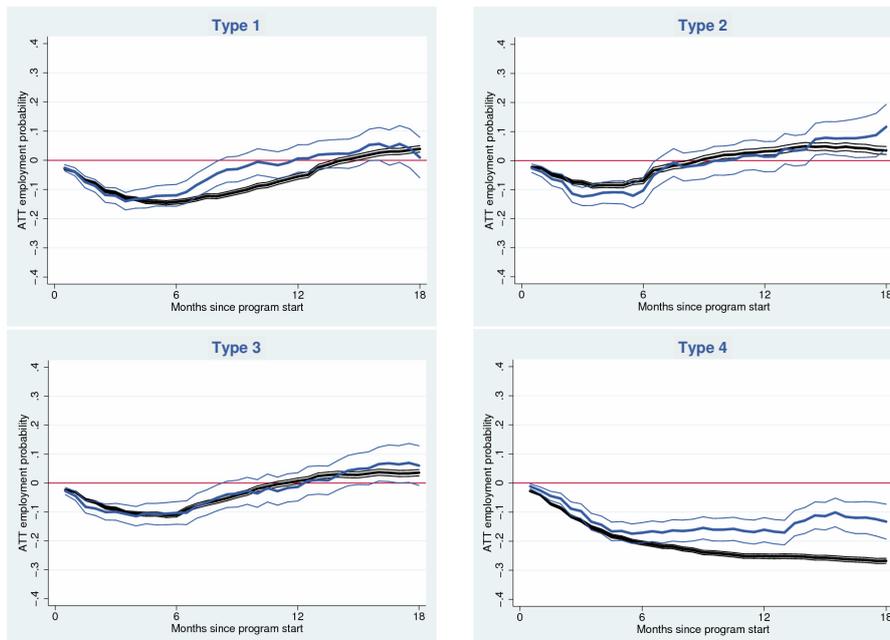
Figure 3: Actual Program Duration by Program Type.



Source: IEB, own calculations.

Note: Kaplan-Meier Estimates. Pre-reform period in black, post-reform period in gray.

Figure 4: Reform Effect (RE).

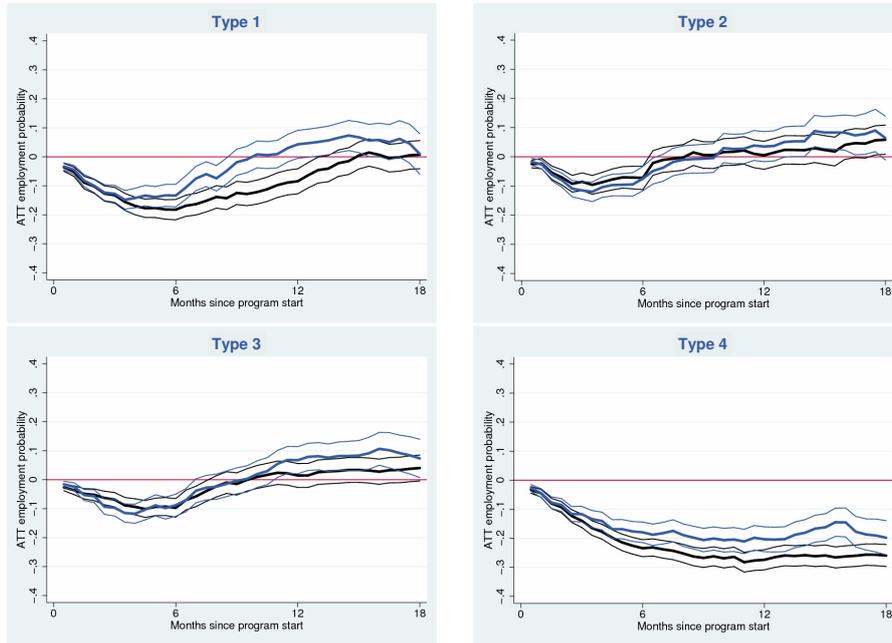


Source: IEB, own calculations.

Note: Pre-reform period in black, post-reform period in blue.

Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 5: Voucher Effect (VE).

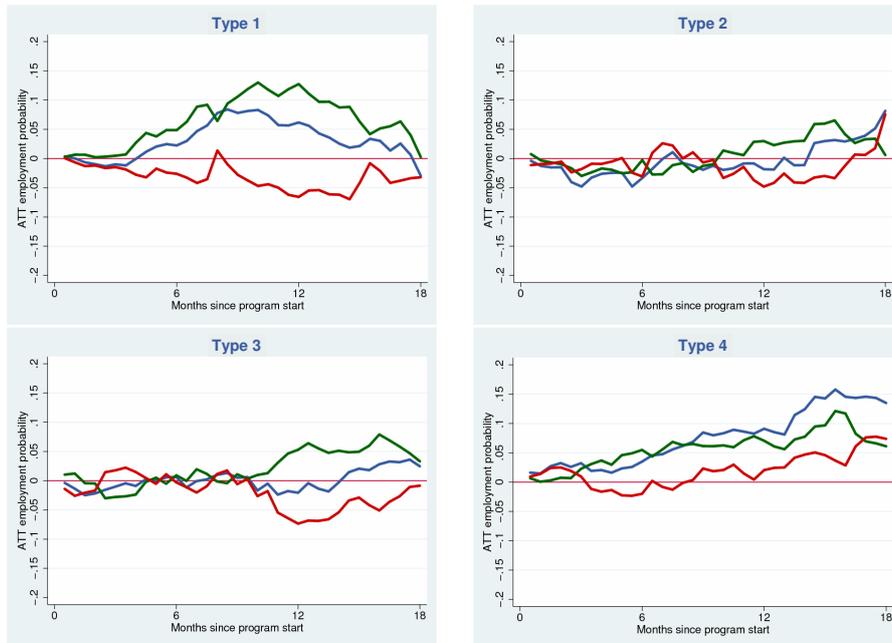


Source: IEB, own calculations.

Note: Pre-reform period in black, post-reform period in blue.

Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 6: Decomposition (Reform Effect vs Voucher Effect vs Selection Effect).



Source: IEB, own calculations.

Note: Total reform effect (RE) in blue, voucher effect (VE) in green, and selection effect (SE) in red.

Table 1: Regression Analyses (OLS).

	Employment after 6 Months	Employment after 12 Months	Employment after 18 Months
<i>Type 1</i>			
Participant	-0.0599 ***	-0.0585 ***	0.1320 ***
After	0.0217 ***	0.0194 ***	-0.0579 ***
Participant \times After	0.0357 ***	0.0591 ***	-0.1194 ***
Vector X_i included	No	No	No
Adjusted R^2	0.0016	0.1085	0.1321
<i>Type 2</i>			
Participant	-0.0009	0.0714 ***	0.0840 ***
After	0.0217 ***	0.0194 ***	-0.0579 ***
Participant \times After	0.0108	0.0386 ***	-0.1024 ***
Vector X_i included	No	No	No
Adjusted R^2	0.0008	0.1087	0.1311
<i>Type 3</i>			
Participant	-0.0452 ***	0.0770 ***	0.1060 ***
After	0.0217 ***	0.0194 ***	-0.0579 ***
Participant \times After	0.0302 ***	0.0496 ***	-0.0798 ***
Vector X_i included	No	No	No
Adjusted R^2	0.0012	0.1086	0.1319
<i>Type 4</i>			
Participant	-0.1436 ***	-0.1636 ***	-0.1597 ***
After	0.0217 ***	0.0194 ***	-0.0579 ***
Participant \times After	0.0282 ***	0.0789 ***	0.0772 ***
Vector X_i included	No	No	No
Adjusted R^2	0.0041	0.1113	0.1312

Source: IEB, own calculations.

Note: *** significant at 1%, ** significant at 5%; * significant at 10%.

Vector X_i contains similar variables as the matching specifications.