# Post-School Human Capital Investments: Training and Lifelong Learning

Dissertation

zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaft
des Fachbereichs Wirtschaftswissenschaft
der Freien Universität Berlin

vorgelegt von

Diplom-Volkswirt Ulf Rinne
geboren am 12. August 1979 in Stadthagen
wohnhaft in Bonn

— eingereicht im Februar 2009 —

*Erstgutachter:* Prof. Dr. Klaus F. Zimmermann

*Zweitgutachter:* Prof. Dr. Ronnie Schöb

*Tag der Disputation:* 14. Juli 2009

Ulf Rinne

# Post-School Human Capital Investments: Training and Lifelong Learning

# Preface

I have written this dissertation while being a Resident Research Affiliate at the Institute for the Study of Labor (IZA). During the past three and a half years I have received support by my supervisor, colleagues, and my family.

First of all I would like to thank my supervisor Klaus F. Zimmermann for his guidance, constant encouragement, and continuous support for my research. I also thank my colleagues at the IZA for their help and support, among others Holger Bonin, Marco Caliendo, Steffen Künn, Andreas Peichl, Ricarda Schmidl, Hilmar Schneider, Marc Schneider, Arne Uhlendorff and Zhong Zhao. The cooperation with them was central for a very friendly and motivating working atmosphere. Moreover, I am grateful to the IZA for financial support through its scholarship program.

Last but not least I would like to thank my parents and my family. Their love and support throughout my life are great sources of strength.

Bonn, February 2009 *Ulf Rinne*

# Contents

# Chapter 1

# Introduction: Post-School Human Capital Investments

*This book studies the effects of post-school human capital investments. Questions which are addressed include the following: Do prime-age skilled unemployed benefit more from training? Can vouchers improve the effectiveness of public training programs? What are the effects of active labor market policy in transition economies? Who engages in lifelong learning, i.e., training activities of employed individuals, and what are the effects of participation?*

# Motivation

*"An investment in knowledge always pays the best interest."* This quotation of Benjamin Franklin (1706–1790) indicates that the importance of investments in human capital was recognized already a long time ago. Nowadays, investments in human capital sometimes even appear to be the "philosopher's stone" to virtually all economic problems and future challenges. For instance, the then President-elect Barack Obama delivered a major address at George Mason University on January 8, 2009 to introduce his *American Recovery and Reinvestment Plan*, an economic stimulus package against the background of the current financial and economic crisis. During his speech he emphasized that part of the plan will be an investment in updating and modernizing American schools: *"To give our children the chance to live out their dreams in a world that is never been more competitive, we will equip tens of thousands of schools, community colleges and public universities with 21st century classrooms, labs and libraries. We will provide new computers, new technology and new training for teachers so that students in Chicago and Boston can compete with children in Beijing for the high-tech, high-wage jobs of the future."* Similar strategies are followed by governments around the world.

In contrast to school education which primarily affects the human capital of future generations, interventions later in the life cycle, e.g., public training programs for the unemployed, potentially have an immediate effect on the current workforce. Therefore, these investments appear attractive from a typically short- or medium-term oriented policymaker's point of view. For example, the government of China decided as a reaction to the global economic crisis to enhance vocational training for the unemployed. A similar approach is part of the second stimulus package initialized in Germany in the beginning of 2009 (*Konjunkturpaket II*). Altogether, these actions indicate that the importance of investments in human capital is broadly appreciated. Based on this observation, this book addresses the underlying question of the effectiveness of these investment decisions in a number of facets.

Before further motivating this book, it seems useful to give a definition of "human capital" and discuss its economic interpretation and effects. In a rather broad definition, human capital refers to the productive capacities of human beings as income producing agents in the economy (Rosen, 1998). Investments in human capital can therefore take various forms which include schooling, training, medical care and migration. To a varying degree these investments have effects on income and consumption, but all of them aim at improving a person's knowledge, skills, qualifications, or health. According to Gary S. Becker, education and training are the most important investments in human capital (Becker, 1962, 1993). In his framework these investments affect a worker's productivity, and thus increase future wages and income. In an alternative scenario education and human capital accumulation mainly serve as signals for otherwise unobservable abilities, importantly without increasing their productivity (Spence, 1973). Although these two opposing approaches towards education and training are not explicitly tested below, it should be kept in mind throughout the lecture of this book that the effects of investments into human capital may work through these two different channels.

This book studies the effects of training and lifelong learning both in a developed country (Germany) and in a transition economy (Serbia). In the latter case, it contributes to the relatively scarce literature on the effects of human capital investments in countries passing through a transitional period. Unemployment rates are high and very persistent in Serbia. This is supposed to be partly an inherited problem and partly due to the prolonged and until 2000 highly irregular transition process. But also after the democratic changes—the fall of the Milošević regime in 2000—unemployment has further increased. In this context it is important to know whether active labor market policy can—at least temporarily—alleviate the unemployment impact of the economic transition process. On the other hand, this book focuses on Germany, a developed and industrialized economy. However, the risk of unemployment is remarkably high among low-skilled and unskilled individuals. Furthermore, the correlation of the level of education and the risk of unemployment has increased in recent years. While around one fifth of workers without vocational degree were unemployed in 2004 and 2005, this share has amounted to merely one out of twenty in the 1970s (Reinberg and Hummel, 2005, 2007). The gap between skill-specific unemployment rates is also relatively large by international standards (OECD, 2006). Another concern is that projections of labor demand and supply show that skill shortages are a problem of increasing importance (Bonin et al., 2007; Schnur and Zika, 2007). In the medium-term future a lack of qualified workers which is accompanied by a high and persistent unemployment rate is considered as a likely scenario for Germany. These developments obviously increase the importance of investments in human capital.

On the other hand, and in particular in the German context, it is frequently argued that the establishment of a low-wage sector is necessary to cope with the high unemployment rates in the lower end of the skill distribution. The current tax and transfer system induces low incentives to work, which is especially a problem for the low-skilled (Zimmermann, 2003). Various proposals have been put forward to increase work incentives, see for example Steiner (2004) and Bonin and Schneider (2006) for discussions. Empirical evidence for the relevance of the incentive problem can be found, e.g., in Schneider and Uhlendorff (2005) and Uhlendorff (2008). Altogether, this line of argumentation stresses the importance of labor supply—and the design of the tax and transfer system—in explaining and overcoming the gap between skill-specific unemployment rates.

But there are moreover developments affecting labor demand which contribute to the increasing importance of investments in human capital on the German labor market. Skill-biased technological (and organizational) is one of these developments (Bauer and Bender, 2004). Berman et al. (1998) argue that skill-biased technological change has shifted demand from less-skilled to skilled workers throughout the developed world. In the United States there is evidence that technical change has been skill-biased probably for most of the twentieth century, but there is further evidence pointing towards an acceleration of the skill bias during the past few decades (Acemoglu, 2002). Spitz-Oener (2006) provides direct evidence for West Germany and shows that changes in skill requirements are indeed similar to those in the United States. Moreover, she finds a sharp increase in nonroutine cognitive tasks in recent decades.

International outsourcing is an additional phenomenon which at least coincides with deteriorating relative wages and employment prospects of low-skilled workers. According to Hummels et al. (2001), international outsourcing grew about 30 percent between 1970 and 1990 in a sample of 14 countries. The labor market consequences of outsourcing are not unambiguous from a theoretical point of view, and thus the question whether workers gain or lose has been analyzed in a number of empirical studies which are summarized, e.g., in Feenstra and Hanson (2001). Geishecker (2008) and Geishecker and Görg (2008) are examples for more recent studies focusing on the impact of international outsourcing on individual job security and wages in Germany, respectively. The picture which arises is twofold: Whereas the former study finds that international outsourcing significantly lowers individual job security similarly across skill groups, the latter study presents evidence for differential wage effects across the skill distribution. More specifically, an increase in international outsourcing reduces wages for low-skilled workers and increases wages for high-skilled workers. Therefore, international outsourcing appears to contribute to increases in the wage gap between skilled and unskilled workers.

Training and lifelong learning are considered to be one solution to the problems and challenges outlined above. This book contributes to the ongoing debate about these issues. It concentrates on post-school human capital investments, and therefore does not study investments into school education or initial training. Instead, training programs for the unemployed are studied, as well as the training receipt of employed individuals after entering the labor market.

# Contribution of this Book

Do prime-age skilled unemployed benefit more from training? Chapter 2 addresses this question as the treatment effects of public training programs for the unemployed in Germany are studied.[1] The picture that has been sketched in previous studies is extended by estimating treatment effects of training programs for different sub-groups of the unemployed with respect to vocational education and age. The results indicate that program participation has a positive impact on employment probabilities for all sub-groups considered. Moreover, participants also seem to find more often higher paid jobs than non-participants. Only little evidence is found in favor of heterogeneous treatment effects, and the magnitude of the differences which are found is quite small. These findings are thus—at least in part—conflicting with the strategy to increasingly provide training to individuals with better employment prospects.

Chapter 3 picks up the topic addressed in the previous chapter.[2] However, its main research question is more specific. The Hartz reform in Germany introduced training vouchers and imposed more selective criteria on the applicants. Although it has been previously shown that the overall impact of the reform on the effectiveness of public training programs was positive, the question remains which features of the reform caused this increase—and to what particular extent. Therefore, besides estimating the total reform effect, the effect induced by changes in the composition of program participants (selection effect) is isolated from the effect based on the introduction of vouchers (voucher effect). The decomposition of the positive reform effect suggests that the selection effect is—if at all—slightly negative, and that the voucher effect increased both, the employment probability and earnings of the participants.

Chapter 4 contributes to the relatively scarce literature on the effect of post-school human capital investments in transition economies.[3] More specifically, it addresses the question how training affects individual outcomes in such an economic environment as the causal impact of participation in an active labor market program—the *Beautiful Serbia* program providing training and temporary work in the construction

[1] Chapter 2 is based on joint work with Marc Schneider and Arne Uhlendorff (Rinne et al., 2007).

[2] Chapter 3 is based on joint work with Arne Uhlendorff and Zhong Zhao (Rinne et al., 2008).

[3] Chapter 4 is based on joint work with Holger Bonin (Bonin and Rinne, 2006).

sector in Serbia—is studied. The program's effectiveness is assessed both in terms of labor market outcomes as well as measures of subjective well-being approximating individual welfare. Interestingly, the positive impact of this particular program appears much stronger judged by subjective well-being than judged by the immediate labor market effect.

Unlike the previous chapters of this book, Chapter 5 turns its focus away from post-school human capital investments of the unemployed. Instead, it complements the preceding chapters as training activities of employed individuals are studied. Using data from the SOEP, two different periods are analyzed: *a*) from 1997 to 2000 and *b*) from 2001 to 2004. While a fairly similar pattern with regard to the incidence of private-sector training in Germany is found in both periods, the picture which arises with respect to the effects of private-sector training on wages is not very robust to the econometric approach. The positive wage effects of about 4–6 percent in both samples in the fixed effects specifications generally decrease quite substantially in the fixed growth rates specifications. With respect to the effect of participation in private-sector training on subsequent employment prospects, the probability of being employed in subsequent years is raised by about 2–3 percentage points in both periods. However, this positive employment effect seems to disappear after around 5 years.

Chapter 6 summarizes the main findings of this book and draws conclusions. Policy implications which can be derived from these results are highlighted. Moreover, potential shortcomings and problems are discussed and an outlook for further research in this area is provided.

The data employed in the empirical analyses in this book come from different sources. While Chapters 2 and 3 are based on administrative data from the Federal Employment Agency (FEA) in Germany, Chapter 4 uses data from a special survey which has been designed for the purpose of evaluating a particular active labor market program in Serbia. Finally, Chapter 5 is based on the German Socio-Economic Panel Study (SOEP) which is a representative longitudinal study of private households starting in 1984.

## Conclusions and Policy Implications

There are several policy implications which can be drawn from the findings presented in this book. The results of Chapter 2 with respect to the effect heterogeneity of public training programs reveal that the effects of these programs are fairly similar across different skill and age groups—taking into account the relative gain compared to the situation without participation. Hence, these findings are conflicting with the strategy to increasingly provide training to individuals with better employment prospects. This strategy has been implemented in Germany as a part of the reform of active labor market policy in 2003. After the reform, caseworkers are asked to evaluate the employment prospects of the unemployed in advance and to provide training only to individuals with a relatively high probability of entering employment after training participation. However, this does not take into account the relative gain compared to the situation without training—which would be important according to the results presented here.

The explicit analysis of the impacts of the labor market reform in 2003 in Chapter 3 confirms the previous results. More specifically, the total reform effect is decomposed into two separate components: The effect which is based on the introduction of vouchers is disentangled from the one which is based on changes in the composition of program participants. The finding that the latter effect plays virtually no role in explaining the overall positive impact of the reform is consistent with the results of the previous chapter. On the other hand, the new allocation process can be regarded as a success since the introduction of vouchers increased the effectiveness of the program under consideration. This result is mainly driven by skilled participants as the reform effect is not significant for the unskilled. While the former group can take advantage from an increased consumer sovereignty, unskilled individuals seem to have problems in adequately using the innovative voucher scheme.

The evaluation of the *Beautiful Serbia* program in Chapter 4 deviates from routine program evaluation by considering subjective measures of individual well-being as possible outcomes. Hence this chapter is linked to the rising economic literature focusing on the concept of happiness as an approximation for the individual welfare scale. The findings concerning the impacts of the program on labor market outcomes suggest that both stages of the program exert a positive influence on the employment prospects of participants. However, the positive effects are not sufficiently strong or clear-cut to be considered statistically significant. On the other hand, significantly positive impacts on a number of dimensions of subjective well-being are found. These results thus provide an example that the positive effects of a policy can appear stronger if it is judged by subjective well-being rather than by labor market effects. The program probably impacted on individual welfare through other channels than the immediate economic status, notably by strengthening self-confidence, job desire and social inclusion of the participants. What can be learned from the evaluation of this particular program in a broader sense is the fact that it appears to be fundamentally important to take into account the demands and requirements of a rigorous evaluation exercise from a very early stage of the program's implementation process. For instance, this includes the design, allocation procedure and data collection.

The preceding analyses are complemented in Chapter 5 by focusing on training activities of employed individuals. While the findings with respect to the wage effects of private-sector training are are not very robust to the econometric approach, there appear to be clearly positive effects on subsequent employment prospects which disappear after around 5 years. Moreover, these positive employment effects seem to be solely based on whether or not an individual engaged in training at all. The respective duration which has been spent in training does not appear to matter in this context. These findings are consistent with the signaling theory (Spence, 1973), but hardly compatible with the theory of human capital (Becker, 1962, 1993).

# Chapter 2

# Do Prime-Age Skilled Unemployed Benefit More from Training? Effect Heterogeneity of Public Training Programs in Germany

*Public training programs are the most important part of active labor market policy in Germany. Although there are already a number of studies analyzing the overall effectiveness of these programs, the empirical evidence on the direction and the extent of potential effect heterogeneity is rather scarce. This chapter fills this gap and investigates the question whether the effects of public training programs are heterogenous with respect to the level of vocational education and age.*[4]

[4]This chapter is based on joint work with Marc Schneider and Arne Uhlendorff (Rinne et al., 2007).

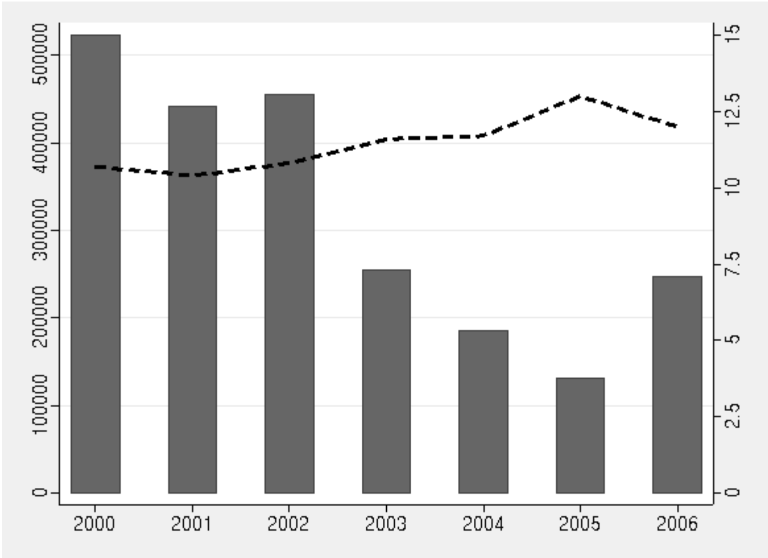
## 2.1 Introduction

One central aim of active labor market policy (ALMP) is to increase the employment prospects of unemployed individuals. For this purpose, the Federal Employment Agency (FEA) in Germany 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 on ALMP measures in 2002 (Eichhorst and Zimmermann, 2007). The most important measures 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 2.1). While more than 500,000 unemployed individuals entered a training program in 2000, this number approached only around 130,000 individuals in 2005 and increased again to nearly 250,000 persons entering such programs in 2006. On the other hand, the unemployment rate remained rather constant during this period.

There already exists a number of studies evaluating the effectiveness of public training programs in Germany. For a recent review of the results see, e.g., Caliendo and Steiner (2005).[5] The results are quite heterogeneous—depending on the method, 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., 2008; Biewen et al., 2007). Hujer et al. (2006) is an example for a recent study which finds negative effects. But these authors concentrate on the duration of the initial unemployment spell, and the negative impact of program participation probably reflects the lock-in effect of training programs. Another example for a recent study reporting negative effects is Lechner and Wunsch (2008). Despite of the difference in the data and the investigation period, 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.

[5]The international literature on the evaluation of ALMP is summarized by Martin and Grubb (2001) and Kluve (2006), among others.

Figure 2.1: Entrants into public training programs and unemployment rate (2000–2006)




*Source:* Federal Employment Agency (FEA).

*Note:* Bars show annual number of entrants into public training programs (left axis); the dashed line represents the average unemployment rate (right axis, in percent).

The above mentioned studies focus on average effects of public training programs, partly differentiated by gender, program type and region. The contribution of this chapter is to extend the picture sketched so far by answering the question whether the effects of public training programs in Germany are heterogenous with respect to the level of vocational education and age.[6] We examine the effects of three types of programs: *a*) programs with a focus on class-room training, *b*) programs with a focus on practical experience, and *c*) training within practice firms, i.e., with a focus on simulating a real working environment. These three types are—in comparison to other ALMP measures in Germany—shorter programs with a median duration between 6 and 8 months.

There does not exist a clear hypothesis for the direction of potential effect heterogeneity. For example, one could think of at least two opposing effects that may affect individuals with and without a vocational degree in a different way. On the

[6] Caliendo et al. (2008) investigate a similar question for job creation schemes in Germany and present evidence for the presence of effect heterogeneity. Although previous results of negative average effects are confirmed in their study, some strata of the population benefit from participation in job creation schemes.

one hand, public training programs may involve diminishing marginal returns, i.e., the more human capital the given individual has already accumulated, the less the training program enhances his or her human capital. On the other hand, the effect of medium-term training programs—the focus of our study—may be positively related to the human capital that has already been accumulated by the individual. In contrast to long-term programs, which are in general aiming to provide a vocational degree, and hence supposedly human capital enhancing by themselves, shorter programs can—at least according to this line of argumentation—only activate already accumulated human capital. In other words, people without a vocational degree would benefit to a smaller extent from participation since skills are provided which are primarily complementary to a vocational degree. The direction and the extent of potential effect heterogeneity is therefore an empirical question and its estimation is the aim of this chapter.

Two recent contributions point into a similar direction as this chapter. Lechner and Wunsch (2008) analyze the effectiveness of several West German training and employment programs from 2000 to 2002. Treatment effects are investigated at a fairly disaggregated level, using a—compared to our study—relatively small inflow sample into unemployment. The authors find evidence for effect heterogeneity and show that job seekers with relatively good *a priori* employment prospects are worse off because of large lock-in effects from which they recover only very slowly, while job seekers with disadvantageous *a priori* employment prospects show below average lock-in effects and positive employment effects for some of the shorter training programs—including job related training. Biewen et al. (2007) use similar data and analyze effect heterogeneity by regressing outcome variables after matching on different socio-economic covariates. They find little heterogeneity along observed characteristics, although in some cases older and less educated participants seem to benefit less or not at all from program participation.

In comparison to Lechner and Wunsch (2008) and Biewen et al. (2007) we have access to a much larger sample of participants in training programs. This allows us to apply matching methods within several sub-groups—e.g., within the sample of women without any vocational degree—and to investigate the effect heterogeneity in greater detail. Moreover, we analyze the effects on monthly earnings by comparing the shares of individuals with and without training in different quartiles of the earnings distribution. This approach provides insights into the effect of program participation on the probability to find higher or lower paid jobs, respectively.

Our analysis is based on an inflow sample into training programs for the year 2002. We ensure that the control group consists of individuals who are as long unemployed as the participants by matching exactly on the previous unemployment duration. Furthermore, a propensity score matching aims to balance differences in a wide range of observable characteristics—including detailed information on previous employment history and regional indicators.

Our results indicate that program participation has a positive impact on employment probabilities for all sub-groups. Participants also seem to find more often higher paid jobs than non-participants. We present only little evidence for the presence of heterogeneous treatment effects and the magnitude of the differences is quite small. If we compare the treatment effects for the most important program type on the employment probability two years after program entry, we find no significant differences with respect to age and vocational education within the same gender. Only if we compare men and women with each other, we find that for this program type young men have a significantly higher treatment effect than older women. Moreover, in case of this program type, the lock-in effect is remarkably shorter for male participants without a vocational degree. Similar results are found for the remaining two program types. The overall picture therefore suggests quite homogenous effects of program participation across sub-groups.

Hence, our results are—at least in part—conflicting with the strategy to increasingly provide training to individuals with better employment prospects. This strategy has been implemented in Germany as a part of the reform of active labor market policy in 2003. After the reform the caseworkers are asked to evaluate the employment prospects of the unemployed in advance and to provide training only to individuals with a relatively high probability of entering employment after training participation. However, this does not take into account the relative gain compared to the situation without training.

The remainder of this chapter is structured as follows: Section 2.2 provides information on our data and briefly describes the program types being analyzed. Section 2.3 presents the econometric methods, and Section 2.4 discusses the results. Finally, Section 2.5 concludes.[7]

## 2.2 Data

We use a sample of a particularly rich administrative data set, the Integrated Employment Biographies (IEB) of the FEA.[8] It contains detailed daily information on employment subject to social security contributions 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.

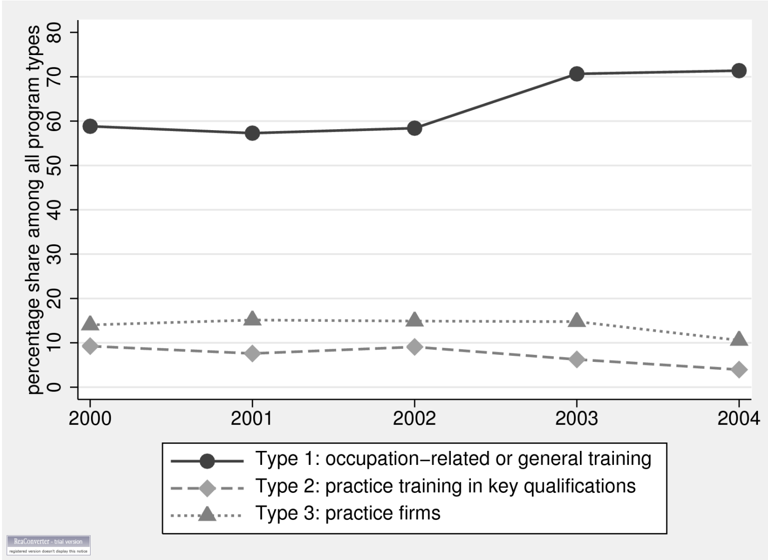
Since the public training programs currently in place in Germany are quite heterogenous, we concentrate on and differentiate between three particular types: *a*) type 1: occupation-related or general training, *b*) type 2: practice training in key qualifications, and *c*) type 3: practice firms. Participants 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'. Often the measure is 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.

Figure 2.2 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

[7]Section 2.6 (Appendix) contains additional figures and tables.

[8]The 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., Jacobebbinghaus and Seth (2007) 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 see, e.g., Schneider et al. (2007).

Figure 2.2: Entrants into public training programs by program type (2000–2004)




*Source:* Federal Employment Agency (FEA).

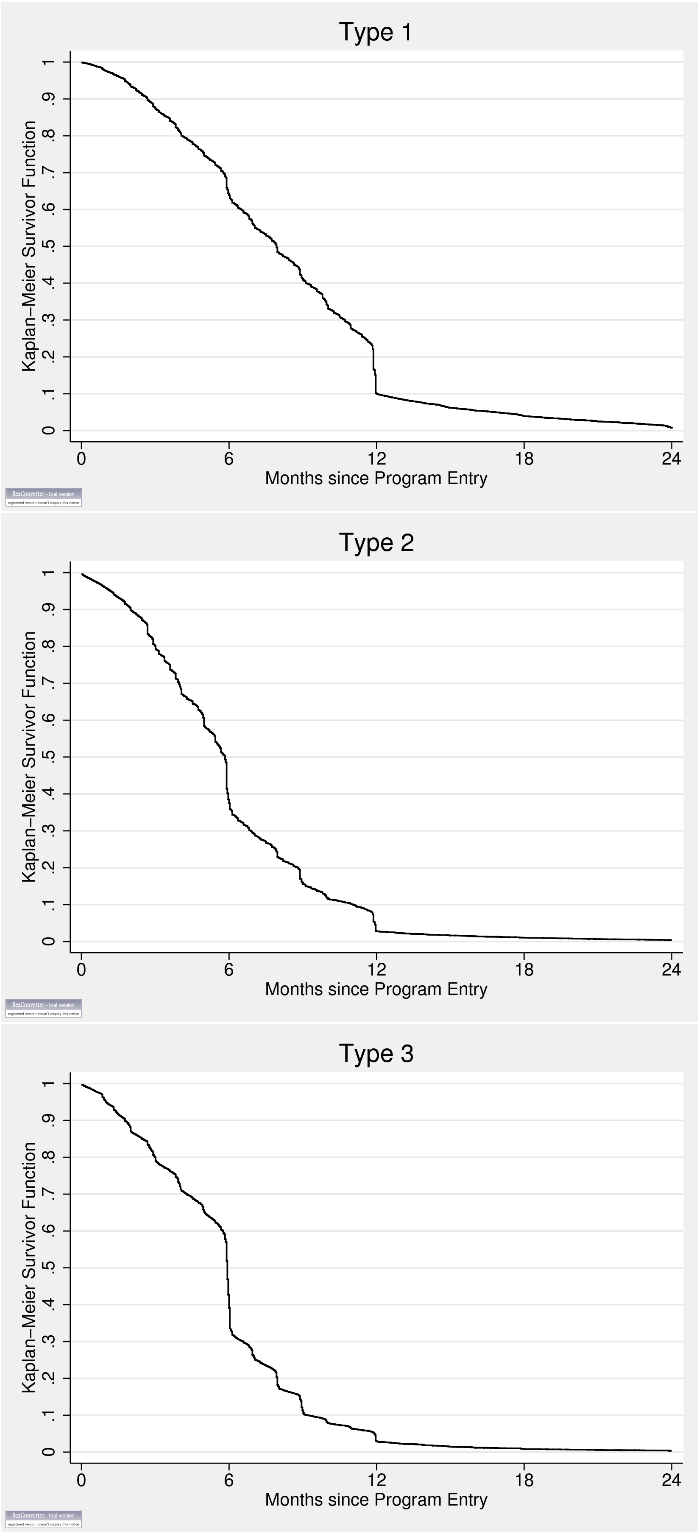
were assigned to this particular type. It became even more important after the reform in 2003 as this share increased to more than 70 percent. Moreover, the three types together account for roughly 85 percent of all participants in public training programs over the period from 2000 to 2004.

Our sample of participants consists of roughly 64,000 unemployed persons entering the three program types in 2002. More precisely, we observe 25,959 participants in type 1, 15,902 participants in type 2, and 22,081 participants in type 3. This sample allows us to draw conclusions on the average participant starting a given program in 2002.[9]

As Figure 2.3 indicates, the three program types are—in comparison to other ALMP measures in Germany—rather shorter measures. After one year, more than 90 percent of the participants have left each type. The median program duration is about 8 months for type 1 and roughly 6 months for types 2 and 3. While a comparatively large fraction of participants finishes type 1 exactly after 12 months, an even larger share finishes type 3 exactly after 6 months. For type 2 we observe a sizeable fraction who ends the measure exactly after 6 or 12 months, respectively.

[9] The number of participants entering a program differs between the analyzed quarters. We take this into account by applying corresponding weights for the calculation of the average treatment effects on the treated.

Figure 2.3: Program duration (2002)




*Source:* IEB, own calculations.

Table 2.1: Sub-sample sizes by vocational education

| | | Male | | Female | |
|---|---|---|---|---|---|
| | | Participants | Non-Participants | Participants | Non-Participants |
| Type 1 | No Degree | 3,206 | 126,383 | 1,756 | 94,621 |
| | Voc. Degree | 11,463 | 208,997 | 9,441 | 173,464 |
| Type 2 | No Degree | 3,510 | 126,383 | 2,602 | 94,621 |
| | Voc. Degree | 5,110 | 208,997 | 4,605 | 173,464 |
| Type 3 | No Degree | 3,932 | 126,383 | 2,061 | 94,621 |
| | Voc. Degree | 8,382 | 208,997 | 7,645 | 173,464 |

*Note:* Completed in-firm training and off-firm training as well as degrees from a vocational school, a technical school, a university, or a university of applied sciences are considered as vocational degrees.

In order to apply the matching approach as described in Section 2.3, around 600,000 non-participants were drawn. Both participants and non-participants are aged between 17 and 65 years.[10]

As we focus on the effect heterogeneity of program participation with respect to vocational education and age, we divide our sample into sub-samples for each program type. With respect to vocational education, the four sub-samples per program type consist of male and female participants and non-participants with and without a vocational degree.[11] Table 2.1 shows that the resulting sample sizes are reasonably large. Only for the sub-sample of female participants in type 1 without a vocational degree we end up with less than 2,000 observations.

With respect to age we divide the sample into six sub-samples for each program type according to gender and three age groups. These age groups were constructed by choosing thresholds in order to end up with sub-samples of more or less the same size. The first age group includes individuals who are 33 years or younger at the (fictitious) program entry, the second group consists of persons aged between 34 and 42 years, and the third group comprises individuals who are at least 43 years old. Here, (fictitious) program entry refers to the point in time where a particular program starts

[10] One could argue for stricter age restrictions, for example because of early retirement regulations in Germany. However, if one is interested in the average effects of treatment on the treated and there are participants older than 55 or 60 years, there is no reason to exclude these individuals.

[11] We consider completed in-firm training and off-firm training as well as degrees from a vocational school, a technical school, a university, or a university of applied sciences as vocational degrees.

Table 2.2: Sub-sample sizes by age group

| | | Male | | Female | |
|---|---|---|---|---|---|
| | | Participants | Non-Participants | Participants | Non-Participants |
| Type 1 | <34 years | 5,759 | 119,615 | 3,582 | 83,993 |
| | 34–42 years | 4,423 | 82,550 | 4,141 | 72,831 |
| | >42 years | 4,536 | 135,350 | 3,509 | 112,446 |
| Type 2 | <34 years | 3,961 | 119,615 | 3,039 | 83,993 |
| | 34–42 years | 1,905 | 82,550 | 1,954 | 72,831 |
| | >42 years | 2,801 | 135,350 | 2,242 | 112,446 |
| Type 3 | <34 years | 4,885 | 119,615 | 2,686 | 83,993 |
| | 34–42 years | 3,483 | 82,550 | 3,231 | 72,831 |
| | >42 years | 3,982 | 135,350 | 3,814 | 112,446 |

for actual participants, while it is used as a reference point for non-participants.[12] The resulting sample sizes are depicted in Table 2.2. While the number of observations of participants is fairly equally distributed within the different sub-samples of program types 1 and 3, this does not entirely apply for type 2. In this case, the groups of male and female participants between 34 and 42 years consist of less than 2,000 observations, respectively.

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 24 months. This period is based on the fact that we focus on program participation in the year 2002, and can observe reliable data for all employment states until December 31, 2004. 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 this outcome measure. Moreover, the administrative data set only includes employment that is subject to social security contributions.[13] Self-employment can thus not be observed in our data. Additionally, we evaluate the effect of program participation on monthly earnings in the primary labor market. We apply the described definition of employment and consider remunerations associated with these spells in terms of monthly earnings.

[12]The specific criteria a non-participant has to meet are further discussed in Section 2.3.

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

## 2.3 Evaluation Approach

Ideally, one would like to compare the outcomes for the individuals participating in public training programs ($Y^1$) with the outcomes for the same individuals if they had not participated ($Y^0$). If $D$ denotes participation in this context—where $D = 1$ if a person participates in the program and $D = 0$ otherwise—the actual outcome for individual $i$ can be written as:

$$Y_i = Y_i^1 \cdot D_i + Y_i^0 \cdot (1 - D_i) \; . \tag{2.1}$$

The individual treatment effect would then be given by the difference $\Delta_i = Y_i^1 - Y_i^0$. However, it is impossible to calculate this difference because one of the outcomes is counterfactual. Instead, the evaluation literature concentrates on population average gains from treatment—usually on the average treatment effect on the treated (ATT or $\Delta_{ATT}$) which is formally given by:

$$\Delta_{ATT} = E(\Delta|D = 1) = E(Y^1|D = 1) - E(Y^0|D = 1) \; . \tag{2.2}$$

It is the principle task of any evaluation study to find a credible estimate for the second term on the right hand side of equation (2.2), which is unobservable.

One possible solution could be to simply compare the mean outcomes of participants and non-participants. However, if $E(Y^0|D = 1) \neq E(Y^0|D = 0)$, estimating the ATT by the difference between the sub-population means of these two groups will yield a selection bias. On the other hand, if treatment assignment is *strongly ignorable*, i.e., if selection is on observable characteristics $X$ (unconfoundedness or conditional independence assumption), and if observable characteristics of participants and non-participants overlap (common support), the matching estimator is an appealing choice to estimate the desired counterfactual (Rosenbaum and Rubin, 1983). Under these conditions, the distribution of the counterfactual outcome $Y^0$ for the participants is the same as the observed distribution of $Y^0$ for the comparison group *conditional on the vector of covariates* $X$. Formally,

$$E(Y^0|X, D = 1) = E(Y^0|X, D = 0) \; . \tag{2.3}$$

Entering this relation into (2.2) allows estimating the ATT by comparing mean outcomes of matched participants and non-participants. Rosenbaum and Rubin (1983) show that if treatment assignment is strongly ignorable *given* $X$, it is also strongly ignorable *given any balancing score that is a function of* $X$.[14] One possible balancing score is the propensity score $P(X)$, i.e., the probability of participating in a given program. Mueser et al. (2007) present evidence that if administrative data is used to measure the performance of training programs, propensity score matching is generally most effective.

There are several propensity score matching methods suggested in the literature, see, e.g., Caliendo and Kopeinig (2008) for an overview. Based on the characteristics of our data, we opt to apply nearest-neighbor matching without replacement. This matching method has the advantage of being the most straightforward matching estimator: a given participant is matched with a non-participant who is closest in terms of the estimated propensity score. We avoid an increased variance of the estimator as we match without replacement (Smith and Todd, 2005a), which is justified since the ratio between participants and non-participants—i.e., potential matching partners—is comparatively high in our data. Hence, the constructed counterfactual outcome is based only on distinct non-participants. To check the sensitivity of our results with respect to the matching algorithm, we additionally applied other methods to our data and find evidence for robust estimates (see Section 2.4.4 for details).

For the variance of the estimated treatment effects, we base our inference on the assumption that the estimators are asymptotically normally distributed. This distribution is derived from the difference of two weighted means of two independent observations. Lechner (2002) employs a similar approach. We checked the accuracy of this approximation by also calculating the variance of the estimated treatment effects based on bootstrapping procedures. Although nearest neighbor matching does not satisfy the basic conditions for the bootstrap and the bootstrap variance diverges from the actual variance (Abadie and Imbens, 2008), this alternative method leads to very similar variances of the estimated treatment effects and does not change the implications presented below.

[14] When there are many covariates, it is impractical to match directly on covariates because of the curse of dimensionality. See, e.g., Zhao (2008) for some comments on this problem.

The focus of the subsequent analysis lies on the differences in treatment effects between separate sub-groups. To assess whether these differences are significantly different from zero, we assume that the treatment effects follow a normal distribution and that they are independent from each other.[15]

The probability of participation in the three program types under consideration is estimated conditional on a number of observable characteristics using binary probit models with participation as the dependent variable. These characteristics include socio-demographic-characteristics (e.g., age, nationality, marital status), regional information (regional type, unemployment rate), educational and vocational attainment, the employment history (four years prior to program entry), and information on the last employment spell (duration, income, business sector).[16] We run these regressions separately for the different sub-samples of participants and non-participants according to program type, gender, and level of vocational education or age group, respectively.

The distribution of the estimated propensity scores is depicted in Figures 2.4 and 2.5. A visual analysis already suggests that the overlap between the group of participants and non-participants is generally sufficient within all sub-samples. Nonetheless, in some cases there are parts of the distribution where participants seem to lack comparable non-participants. However, by using the usual 'Minmax' criterion, where treated individuals are excluded from the sample whose propensity score lies above the highest propensity score in the comparison group, only 4 (24) individuals are dropped in the sub-samples previously stratified with respect to the level of vocational eduction (with respect to the age groups).

After estimating the propensity score we match each participant with a distinct non-participant within the different sub-samples by exact covariate matching plus propensity score matching.[17] Non-participants are required to not having participated in the *respective* type of public training program before and in the quarter of the participant's program entry. The variables used for exact matching are previous duration of unemployment (in months) and quarter of (fictitious) program entry. Therefore, we

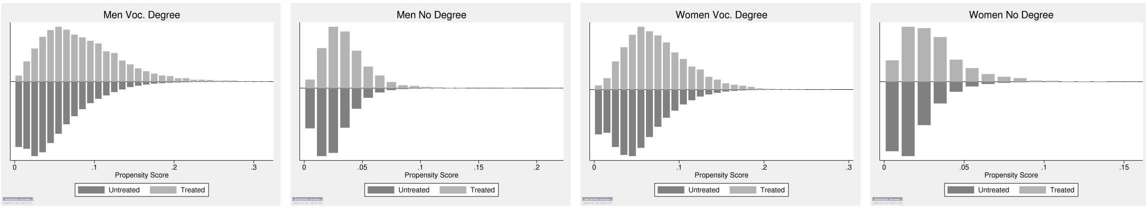
[15]If we drop the assumption of independence, i.e., if we allow for non-zero correlation between treatment effects, implications only marginally change.

[16]The exact specifications are not reported here, but available upon request.

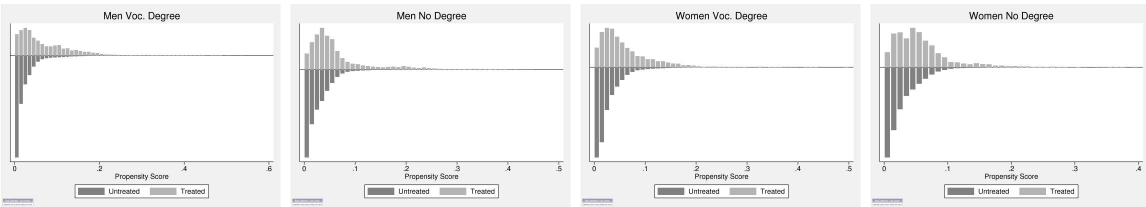
[17]The matching algorithm is implemented using the `PSMATCH2` Stata ado-package by Leuven and Sianesi (2003).

Figure 2.4: Common support by vocational education

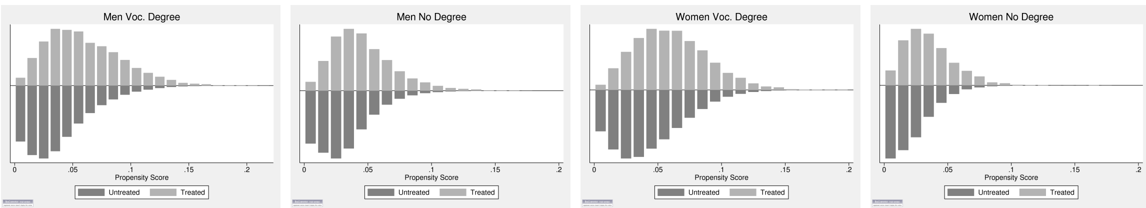
(a) Type 1




(b) Type 2



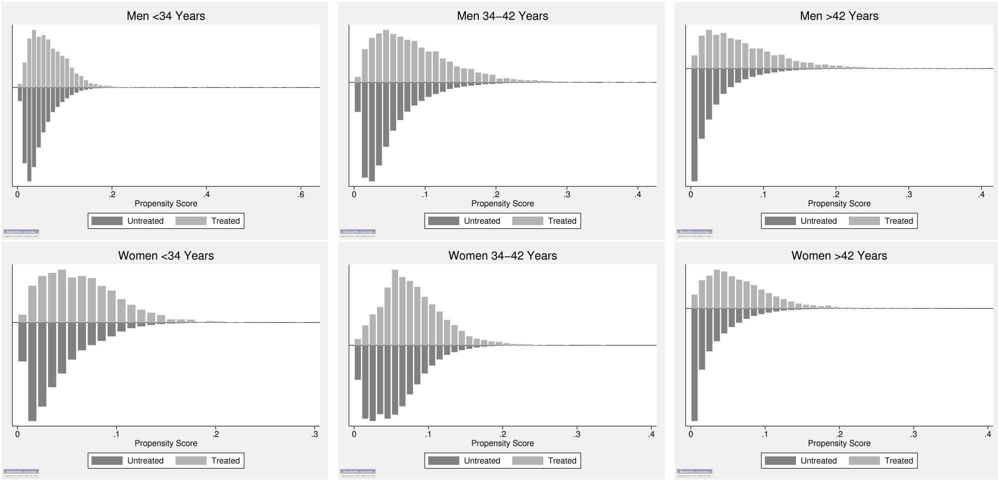

(c) Type 3



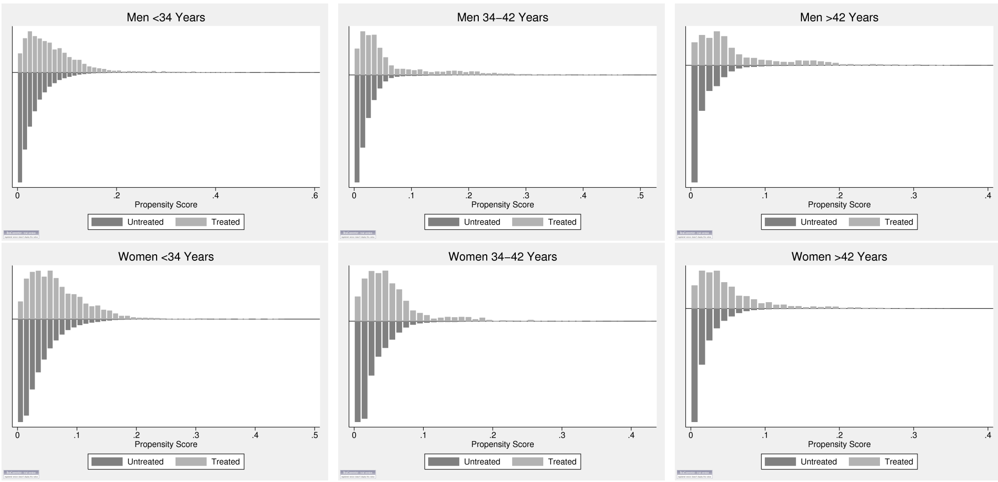

*Note:* Distribution of the estimated propensity scores before matching. Participants are depicted in the upper half, non-participants in the lower half of each figure.

Figure 2.5: Common support by age group

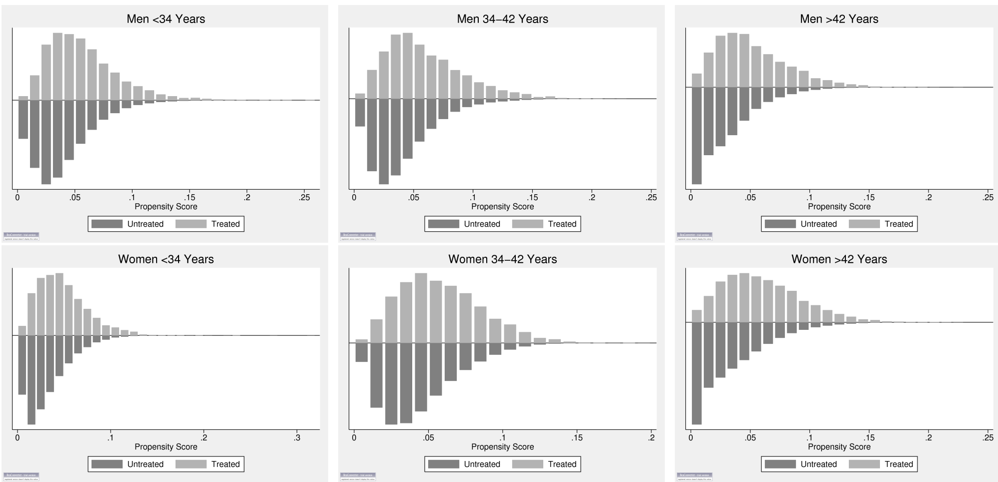
(a) Type 1




(b) Type 2




(c) Type 3




*Note:* Distribution of the estimated propensity scores before matching. Participants are depicted in the upper half, non-participants in the lower half of each figure.

stratify the sub-samples by these variables first, and then implement propensity score matching for each cell without replacing the matched non-participant.

This procedure ensures that matched participants and non-participants are (a) previously unemployed for the same duration at the (fictitious) program entry, and (b) (fictitiously) entering the program in the same quarter. While the latter condition makes sure that seasonal influences are constant and that the observation period is the same for matched pairs, the former condition builds on similar arguments as, e.g., Sianesi (2004) put forward. She argues that participation decisions in ALMP are to be viewed subsequently over time in unemployment, since choices faced by unemployed individuals are not whether to participate or not to participate at all, but rather whether to join a program now or not to participate for now. According to this line of argumentation, it is fundamental to ensure the same elapsed unemployment duration for matched treated and controls.

However, we use program entry as our point of reference rather than following entrants into unemployment over time (inflow sample into unemployment). The estimates we present below can thus be viewed as the outcome of the joining/waiting-decision after the same elapsed duration of unemployment for given individuals. Our approach allows us to estimate the ATT for average participants in given program types in 2002—as opposed to the ATT for participants in given program types of a specific entry cohort into unemployment. Importantly, exact matching on the previous unemployment duration only considers the past up to the (fictitious) entry into the given program. Future outcomes are not considered in this context. In particular, non-participants can potentially participate in the given program type *after* the (fictitious) program entry. Sianesi (2004) employs a similar definition of non-participation. She argues—for the case of Sweden—that in principle any unemployed individual will join a program at some time, provided he remains unemployed long enough. We think that Sweden is similar to Germany in this respect. Hence, a restriction on future outcomes—i.e., to require non-participation in the follow-up period after the (fictitious) program entry—affects estimated treatment effects negatively (Stephan, 2008) since a substantial fraction of the 'never treated'-individuals *de facto* leaves the unemployment register.[18]

After forming the matched pairs, a suitable way to assess the matching quality is comparing the standardized bias before matching, $SB^b$, to the standardized bias after matching, $SB^a$. The standardized biases are defined as

$$SB^b = \frac{(\overline{X}_1 - \overline{X}_0)}{\sqrt{0.5 \cdot (V_1(X) + V_0(X))}} \quad ; \quad SB^a = \frac{(\overline{X}_{1M} - \overline{X}_{0M})}{\sqrt{0.5 \cdot (V_{1M}(X) + V_{0M}(X))}} \ , \tag{2.4}$$

where $X_1$ ($V_1$) is the mean (variance) in the treated group before matching and $X_0$ ($V_0$) the analogue for the comparison group. $X_{1M}$ ($V_{1M}$) and $X_{0M}$ ($V_{0M}$) are the corresponding values after matching (Rosenbaum and Rubin, 1985). Following the example of Sianesi (2004) we also re-estimate the propensity score on the matched sample to compute the pseudo-$R^2$ before and after matching.

Tables 2.3 and 2.4 suggest that the quality of our matching procedures is satisfactory: the percentage biases of a number of covariates are apparently reduced and any significant differences in these covariates disappear after matching. More specifically, the standardized bias for each covariate is below 6 percent after matching. Moreover, the mean standardized bias of the matched samples are noticeably smaller than that of the unmatched sample (between 0.8 and 1.9 percent in the different sub-samples). Likewise, the pseudo-$R^2$ after matching are fairly low and decrease substantially compared to before matching. Tables 2.9–2.11 (see Appendix) include more details concerning the matching quality by program type, e.g., regarding the balancing of covariates.

Training programs may have an influence on the employment probability as well as on the (potential) earnings of the participants. Evaluating the causal effect on the employment probability is straightforward and given by a simple comparison of treatment and control group. In contrast to that, a simple comparison of the realized wages does not give us a clear measure of the causal effect of program participation. Realized earnings are the product of the employment probability and the observed individual earnings, i.e., realized earnings are only a 'crude' measure of the effect on

[18]For instance, Lechner and Wunsch (2008) require non-participation in the follow-up period after the (fictitious) program entry for comparison individuals. Applying the same definition of non-participation to our data lowers the estimated treatment effects (see Section 2.4.4 for details). 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.

Table 2.3: Matching quality within sub-samples by vocational education

| Type | Sex | Vocational Education | | Mean %-Bias | Max. %-Bias | Pseudo-$R^2$ |
|---|---|---|---|---|---|---|
| 1 | Female | Voc. Degree | Before Matching | 8.905 | 31.803 | 0.041 |
| | | | After Matching | 0.972 | 2.611 | 0.001 |
| 1 | Male | Voc. Degree | Before Matching | 12.398 | 40.073 | 0.063 |
| | | | After Matching | 0.983 | 2.837 | 0.001 |
| 1 | Female | No Degree | Before Matching | 10.465 | 26.845 | 0.036 |
| | | | After Matching | 1.330 | 3.597 | 0.002 |
| 1 | Male | No Degree | Before Matching | 9.541 | 22.733 | 0.030 |
| | | | After Matching | 1.289 | 2.971 | 0.002 |
| 2 | Female | Voc. Degree | Before Matching | 10.900 | 46.361 | 0.069 |
| | | | After Matching | 1.157 | 3.355 | 0.002 |
| 2 | Male | Voc. Degree | Before Matching | 11.909 | 47.814 | 0.101 |
| | | | After Matching | 1.153 | 3.442 | 0.002 |
| 2 | Female | No Degree | Before Matching | 9.682 | 40.362 | 0.058 |
| | | | After Matching | 1.534 | 3.408 | 0.003 |
| 2 | Male | No Degree | Before Matching | 9.868 | 43.488 | 0.075 |
| | | | After Matching | 1.098 | 3.578 | 0.001 |
| 3 | Female | Voc. Degree | Before Matching | 8.401 | 26.462 | 0.026 |
| | | | After Matching | 0.988 | 3.403 | 0.001 |
| 3 | Male | Voc. Degree | Before Matching | 9.833 | 39.831 | 0.044 |
| | | | After Matching | 1.030 | 2.998 | 0.001 |
| 3 | Female | No Degree | Before Matching | 7.849 | 21.602 | 0.023 |
| | | | After Matching | 1.465 | 4.192 | 0.003 |
| 3 | Male | No Degree | Before Matching | 10.084 | 24.600 | 0.033 |
| | | | After Matching | 1.202 | 3.847 | 0.002 |

*Note:* Reported indicators refer to 75 variables that are at least included in the specification.

Table 2.4: Matching quality within sub-samples by age group

| Type | Sex | Age | | Mean %-Bias | Max. %-Bias | Pseudo-$R^2$ |
|---|---|---|---|---|---|---|
| 1 | Female | <34 years | Before Matching | 12.391 | 48.062 | 0.055 |
| | | | After Matching | 0.949 | 3.672 | 0.002 |
| 1 | Male | <34 years | Before Matching | 11.567 | 40.533 | 0.047 |
| | | | After Matching | 0.746 | 2.222 | 0.001 |
| 1 | Female | 34–42 years | Before Matching | 10.233 | 41.393 | 0.049 |
| | | | After Matching | 1.141 | 3.452 | 0.001 |
| 1 | Male | 34–42 years | Before Matching | 14.809 | 35.076 | 0.067 |
| | | | After Matching | 1.161 | 3.511 | 0.002 |
| 1 | Female | >42 years | Before Matching | 16.253 | 65.459 | 0.097 |
| | | | After Matching | 1.288 | 3.836 | 0.002 |
| 1 | Male | >42 years | Before Matching | 17.296 | 60.275 | 0.104 |
| | | | After Matching | 0.881 | 2.700 | 0.001 |
| 2 | Female | <34 years | Before Matching | 11.615 | 51.886 | 0.070 |
| | | | After Matching | 1.539 | 4.256 | 0.003 |
| 2 | Male | <34 years | Before Matching | 12.085 | 48.487 | 0.092 |
| | | | After Matching | 1.529 | 4.635 | 0.003 |
| 2 | Female | 34–42 years | Before Matching | 9.746 | 42.853 | 0.052 |
| | | | After Matching | 1.736 | 5.497 | 0.005 |
| 2 | Male | 34–42 years | Before Matching | 12.672 | 58.718 | 0.094 |
| | | | After Matching | 1.641 | 4.787 | 0.004 |
| 2 | Female | >42 years | Before Matching | 10.395 | 53.757 | 0.077 |
| | | | After Matching | 1.280 | 4.519 | 0.003 |
| 2 | Male | >42 years | Before Matching | 12.402 | 58.541 | 0.102 |
| | | | After Matching | 1.887 | 5.680 | 0.003 |
| 3 | Female | <34 years | Before Matching | 11.860 | 36.372 | 0.042 |
| | | | After Matching | 1.491 | 5.067 | 0.003 |
| 3 | Male | <34 years | Before Matching | 9.743 | 30.744 | 0.039 |
| | | | After Matching | 0.998 | 3.283 | 0.001 |
| 3 | Female | 34–42 years | Before Matching | 8.642 | 30.207 | 0.025 |
| | | | After Matching | 0.842 | 2.871 | 0.001 |
| 3 | Male | 34–42 years | Before Matching | 9.124 | 29.356 | 0.036 |
| | | | After Matching | 1.323 | 3.596 | 0.002 |
| 3 | Female | >42 years | Before Matching | 12.351 | 56.536 | 0.062 |
| | | | After Matching | 1.414 | 3.375 | 0.002 |
| 3 | Male | >42 years | Before Matching | 12.472 | 58.976 | 0.064 |
| | | | After Matching | 0.991 | 3.276 | 0.002 |

*Note:* Reported indicators refer to 75 variables that are at least included in the specification.

productivity (Lechner and Melly, 2007). Measuring the causal effect on the earnings would require taking into account the selection into the observed employment, e.g., by making use of an instrument which influences the employment probability but not the earnings.[19] In general, such an instrument is not available.

We argue that we can nonetheless gain interesting insights into the effects of participation on the (observed) monthly earnings by comparing the earnings distributions between treated and controls. From a policy point of view, it is interesting to know to which extent the share of individuals ending up in higher paid jobs is increased by participating in training programs. This effect is given by a comparison of the shares of individuals entering a job above certain thresholds or within a given strata. This is not the causal effect on the—only partially observed—earnings capacity, but the causal effect on the realized monthly earnings. And in contrast to a simple comparison of mean earnings, we can gather information on whether new jobs are mainly lower or higher paid jobs—given participation or non-participation. The mentioned thresholds (or strata) are in our case based on the overall distribution of monthly earnings two years after program entry. In other words, we calculate quartiles of the earnings distribution for participants and matched non-participants across program types—given positive monthly earnings are observed—and compare the fraction of treated and controls within these thresholds for the sub-groups under consideration.

## 2.4 Results

After applying the matching approach as described above, the ATT can be calculated as the difference in mean outcomes between the groups of matched participants and non-participants. Below we present estimates of differences in employment probabilities and monthly earnings from employment in the primary labor market for a period of two years after the (fictitious) program entry.[20] While average treatment effects for the whole sample are discussed in Subsection 2.4.1, the effect heterogeneity of these

[19]Lechner and Melly (2007) propose to estimate bounds for the earnings effects as an alternative method. However, this approach goes beyond the scope of this chapter.

[20]We 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).

effects with respect to vocational education is regarded in Subsection 2.4.2 and with respect to age in Subsection 2.4.3. Subsequently, we assess the sensitivity of our results in Subsection 2.4.4.

### 2.4.1 Average Treatment Effects

To obtain a general impression of the ATT on employment probabilities and monthly earnings, we aggregate the matched sub-groups for each program type and calculate treatment effects as the difference in mean outcomes between participants and non-participants in the resulting samples. Although this procedure was implemented both for the matched sub-samples previously stratified according to the level of vocational education and with respect to age groups, the latter results are not reported in this section since they do not differ.[21]

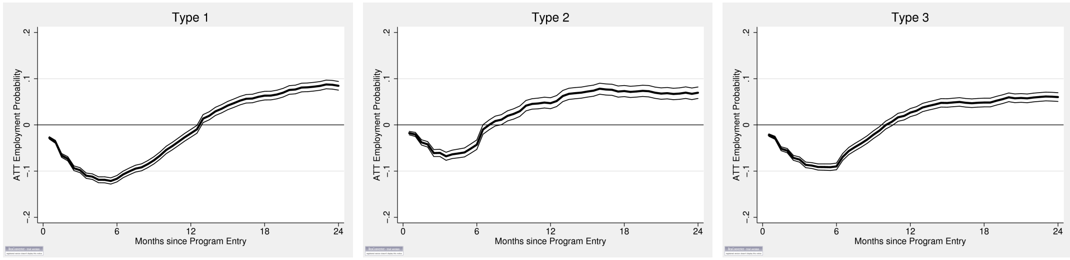
The treatment effects display ATT on employment probabilities and monthly earnings effects, respectively, for a period of 24 months after the (fictitious) program entry. These effects, for one thing, consist of lock-in effects for the group of participants due to reduced search activities while participating in a program (van Ours, 2004), and for another (as an opposing effect), of an expected increase in employment probabilities through and after completing the program.

**Employment Probabilities** For program type 1, we find that participation has a significantly positive impact on the probability of being employed starting about 13 months after program entry (see Figure 2.6). However, in previous months the impact of being locked-in in the program leads to significantly negative point estimates of the ATT. Two years after program entry we observe a point estimate of about 8.5 percentage points.

Our findings on the general effectiveness of type 2 are also rather positive. Although the effect of being locked-in in the program is apparent, we find that participation (significantly) increases the probability of being employed already starting about 7 (8) months after program entry. Two years after program entry, the point estimate is slightly lower than for type 1, but still amounts to roughly 7.5 percentage points.

[21]However, these results are available from the authors upon request.

Figure 2.6: ATT employment probabilities




*Note:* Thick lines are point estimates of the ATT based on aggregated matched sub-samples with respect to vocational education, while thin lines represent 95 percent confidence intervals. The ATT for for the aggregated matched sub-samples with respect to age look very similar and are thus not displayed.

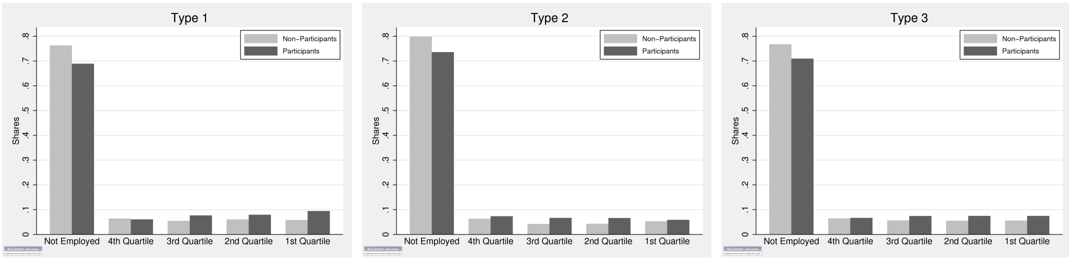
A positive impact of participation on employment probabilities is also found for program type 3. We compute a point estimate of about 6 percentage points two years after program entry. Here, the treatment effect becomes significantly positive about 10 months after entering the program.

**Earnings Effects** For all program types and over the whole two-year-period after program entry, the ATT on monthly earnings (see Figure 2.12, Appendix) do not exhibit major differences compared to the ATT on employment probabilities described above and will thus not be further discussed in Subsections 2.4.2 and 2.4.3.[22] However, to give an idea about the magnitude of the monthly earnings effects two years after entering the program, participants in type 1 (2 and 3) earn about 130 Euros (100 Euros) per month more than comparable non-participants.

Figure 2.7 displays the monthly earnings distribution along with the employment effects two years after program entry. Again the above described positive employment effects for each program type can be observed. Moreover, it is possible to assess to what type of jobs (in terms of monthly earnings) the positive employment effects lead. Therefore, individuals with earnings from unsubsidized employment in the primary labor market are divided into quartiles. The fourth quartile includes gross monthly earnings up to 1,050 Euros, the third quartile up to 1,439 Euros, the second up to 1,890 Euros and the first quartile includes monthly earnings above 1,890 Euros. The graphs show that participants of types 1 and 3 enter additional jobs in the top three

[22]Figures and Tables concerning these results can be found in Section 2.6 (Appendix).

Figure 2.7: Earnings distribution and employment effects 24 months after program entry




*Note:* Quartiles are based on the distribution of monthly earnings in the matched samples, aggregated across program types. 4th quartile: gross monthly earnings <1,050 Euros; 3rd quartile: gross monthly earnings 1,050–1,439 Euros; 2nd quartile: gross monthly earnings 1,440–1,890 Euros; 1st quartile: gross monthly earnings >1,890 Euros.

quartiles of the earnings distribution (and in particular show a significant increase in the first quartile), while for participants of type 2 we observe significantly increased shares in both the middle quartiles. For all types, the fraction of participants in the bottom quartile is about the same as without participation.

### 2.4.2 Treatment Effects: Vocational Education

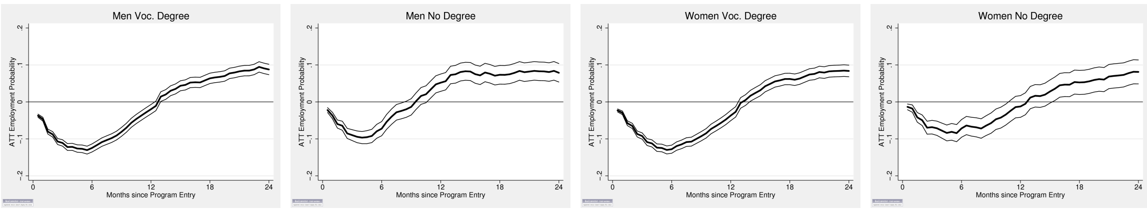
The following section describes the effect heterogeneity of treatment effects with respect to the participants' level of vocational education. For this purpose, we distinguish between male and female participants with and without a vocational degree.

**Employment Probabilities** For all three types, the resulting treatment effects across all sub-groups are quite similar within each type. Nonetheless, some differences appear between the three types (as already discussed in the previous section) and some small differences occur within each type.

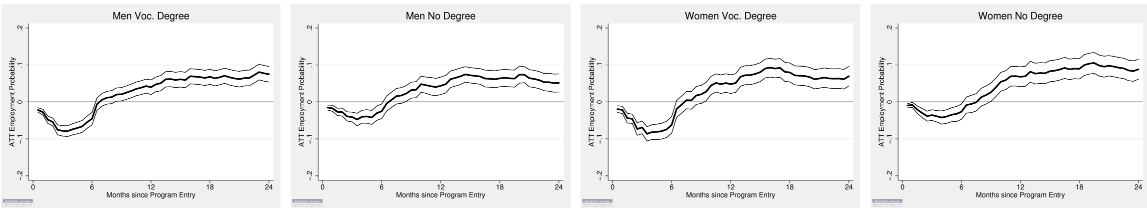
For type 1, positive treatment effects can be observed starting about 13 months after program entry for individuals with a vocational degree (see Figure 2.8a). For individuals without a vocational degree, we find (significantly) positive treatment effects after 9 (10) for men, and (significantly) positive treatment effects after 13 (16) months for women. Also the point estimates of the ATT two years after entering the program show minor differences for the different sub-groups: for male (female)

Figure 2.8: ATT employment probabilities by vocational education

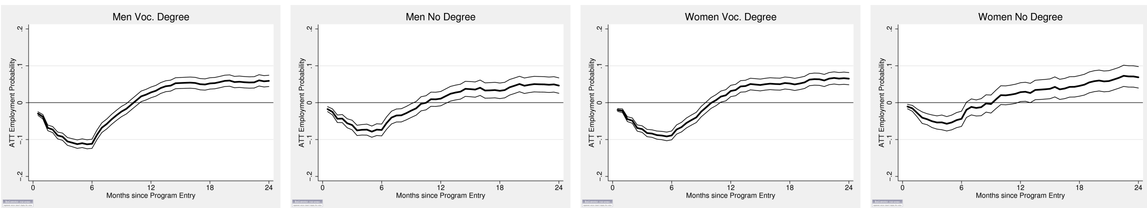
(a) Type 1




(b) Type 2




(c) Type 3




*Note:* Thick lines are point estimates of the ATT for the respective sub-group, while thin lines represent 95 percent confidence intervals.

individuals with a vocational degree they amount to about 9 (8.5) percentage points, while for both sub-groups of individuals without a vocational degree they amount to roughly 8 percentage points (compare Table 2.5).

For type 2, treatment effects become positive about 6–7 months after program entry for all sub-groups (see Figure 2.8b). However, the magnitude of the estimated treatment effects varies (compare Table 2.5): while the point estimate for female participants without a vocational degree amounts to almost 9 percentage points two years after entering the program, the ATT is estimated to be 5 percentage points for male participants without a vocational degree. With roughly 7.5 percentage points the estimated ATT for individuals with a vocational degree lie in between.

For type 3, we observe a similar pattern for women, irrespective of their level of vocational education (see Figure 2.8c). The estimated ATT become positive about

Table 2.5: ATT employment probabilities by vocational education

| Type | Sex | Vocational Education | Month After Program Entry | Emp. Prob. NP | Emp. Prob. P | $\Delta_{ATT}$ | Lower Bound 95% CI | Upper Bound 95% CI |
|---|---|---|---|---|---|---|---|---|
| 1 | Female | Voc. Degree | 6 | 0.1905 | 0.0625 | -0.1280 | -0.1390 | -0.1171 |
| | | | 12 | 0.2370 | 0.2001 | -0.0368 | -0.0507 | -0.0230 |
| | | | 18 | 0.2709 | 0.3330 | 0.0621 | 0.0466 | 0.0776 |
| | | | 24 | 0.2812 | 0.3650 | 0.0838 | 0.0681 | 0.0996 |
| 1 | Male | Voc. Degree | 6 | 0.2304 | 0.1058 | -0.1246 | -0.1354 | -0.1137 |
| | | | 12 | 0.2677 | 0.2479 | -0.0198 | -0.0325 | -0.0070 |
| | | | 18 | 0.3143 | 0.3781 | 0.0638 | 0.0498 | 0.0778 |
| | | | 24 | 0.3144 | 0.4022 | 0.0878 | 0.0738 | 0.1019 |
| 1 | Female | No Degree | 6 | 0.1392 | 0.0675 | -0.0717 | -0.0956 | -0.0478 |
| | | | 12 | 0.1758 | 0.1635 | -0.0123 | -0.0417 | 0.0171 |
| | | | 18 | 0.1954 | 0.2482 | 0.0527 | 0.0201 | 0.0853 |
| | | | 24 | 0.1805 | 0.2618 | 0.0812 | 0.0488 | 0.1136 |
| 1 | Male | No Degree | 6 | 0.1648 | 0.0925 | -0.0724 | -0.0916 | -0.0532 |
| | | | 12 | 0.1958 | 0.2479 | 0.0521 | 0.0287 | 0.0755 |
| | | | 18 | 0.2373 | 0.3107 | 0.0734 | 0.0481 | 0.0987 |
| | | | 24 | 0.2286 | 0.3075 | 0.0788 | 0.0539 | 0.1038 |
| 2 | Female | Voc. Degree | 6 | 0.1928 | 0.1307 | -0.0621 | -0.0848 | -0.0395 |
| | | | 12 | 0.2226 | 0.2706 | 0.0481 | 0.0229 | 0.0733 |
| | | | 18 | 0.2402 | 0.3199 | 0.0797 | 0.0541 | 0.1054 |
| | | | 24 | 0.2530 | 0.3227 | 0.0696 | 0.0433 | 0.0960 |
| 2 | Male | Voc. Degree | 6 | 0.1944 | 0.1489 | -0.0455 | -0.0624 | -0.0286 |
| | | | 12 | 0.2415 | 0.2813 | 0.0398 | 0.0201 | 0.0595 |
| | | | 18 | 0.2634 | 0.3316 | 0.0682 | 0.0476 | 0.0887 |
| | | | 24 | 0.2603 | 0.3352 | 0.0749 | 0.0539 | 0.0960 |
| 2 | Female | No Degree | 6 | 0.1224 | 0.0961 | -0.0263 | -0.0471 | -0.0055 |
| | | | 12 | 0.1494 | 0.2174 | 0.0680 | 0.0412 | 0.0949 |
| | | | 18 | 0.1606 | 0.2502 | 0.0896 | 0.0621 | 0.1171 |
| | | | 24 | 0.1614 | 0.2497 | 0.0883 | 0.0616 | 0.1150 |
| 2 | Male | No Degree | 6 | 0.1458 | 0.1205 | -0.0253 | -0.0450 | -0.0056 |
| | | | 12 | 0.1743 | 0.2180 | 0.0436 | 0.0200 | 0.0673 |
| | | | 18 | 0.1867 | 0.2505 | 0.0637 | 0.0409 | 0.0865 |
| | | | 24 | 0.2038 | 0.2552 | 0.0514 | 0.0269 | 0.0758 |
| 3 | Female | Voc. Degree | 6 | 0.1767 | 0.0879 | -0.0888 | -0.1007 | -0.0768 |
| | | | 12 | 0.2222 | 0.2540 | 0.0317 | 0.0164 | 0.0470 |
| | | | 18 | 0.2637 | 0.3158 | 0.0521 | 0.0357 | 0.0685 |
| | | | 24 | 0.2724 | 0.3376 | 0.0652 | 0.0486 | 0.0817 |
| 3 | Male | Voc. Degree | 6 | 0.2352 | 0.1237 | -0.1115 | -0.1239 | -0.0991 |
| | | | 12 | 0.2644 | 0.2923 | 0.0279 | 0.0134 | 0.0424 |
| | | | 18 | 0.3106 | 0.3628 | 0.0523 | 0.0369 | 0.0676 |
| | | | 24 | 0.3093 | 0.3684 | 0.0591 | 0.0438 | 0.0745 |
| 3 | Female | No Degree | 6 | 0.1113 | 0.0675 | -0.0437 | -0.0644 | -0.0231 |
| | | | 12 | 0.1528 | 0.1814 | 0.0286 | 0.0031 | 0.0541 |
| | | | 18 | 0.1826 | 0.2282 | 0.0456 | 0.0176 | 0.0735 |
| | | | 24 | 0.1844 | 0.2533 | 0.0688 | 0.0398 | 0.0979 |
| 3 | Male | No Degree | 6 | 0.1637 | 0.0896 | -0.0740 | -0.0900 | -0.0580 |
| | | | 12 | 0.1991 | 0.2101 | 0.0110 | -0.0084 | 0.0304 |
| | | | 18 | 0.2344 | 0.2666 | 0.0322 | 0.0111 | 0.0533 |
| | | | 24 | 0.2302 | 0.2765 | 0.0464 | 0.0255 | 0.0673 |

*Note:* NP: Non-Participants; P: Participants; CI: confidence interval.

10 months after program entry and lie between 6.5 and 7 percentage points two years after program entry (compare Table 2.5). While it also takes about 10 months after program entry to observe positive treatment effects for male participants of this program type, the point estimates two years after program entry are lower than for female participants. For male participants with a vocational degree the point estimate of the ATT amounts to about 6 percentage points, while it only lies around 4.5 percentage points for men without a vocational degree.

In summary, it is important to note that we find—with respect to the ATT two years after program entry—only one significant difference between sub-groups. For program type 2, men without a degree gain significantly less by participating in training than women without a degree. No other significant differences are observed.

**Earnings Effects** There are also only minor differences in monthly earnings effects across sub-groups within each program type, as is the case for the ATT on employment probabilities just discussed.

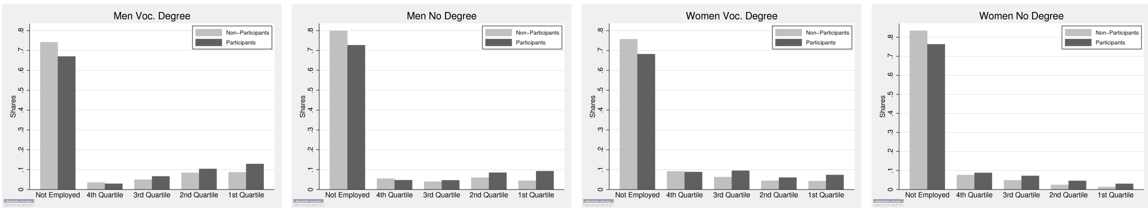
For type 1, Figure 2.9a shows the monthly earnings distribution along with the employment effects two years after program entry. Across all sub-groups, an additional fraction of participants enters jobs in the top quartiles of the earnings distribution—especially in the first quartile, where this increase is significantly positive for all sub-groups. Furthermore, the share of participants in the bottom quartile of the earnings distribution is at most equal or even significantly lower (for male participants with a vocational degree) compared to matched non-participants.

When looking at the monthly earnings distribution for type 2 in company with the employment effects two years after program entry (see Figure 2.9b), we can distinguish a slightly different impact of program participation for men and women: while we observe employment in additional jobs located in the second and third quartile of the earnings distribution for men (with a tendency towards the top quartile for those with a vocational degree), we find that additional jobs are mainly located in the second and third quartile with a tendency towards the bottom quartile for women—especially for those without a vocational degree.

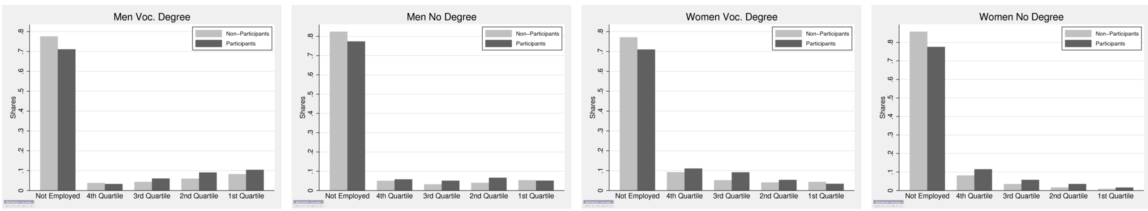
For type 3, again, we find only minor differences across the sub-groups (see

Figure 2.9: Earnings distribution and employment effects 24 months after program entry by vocational education

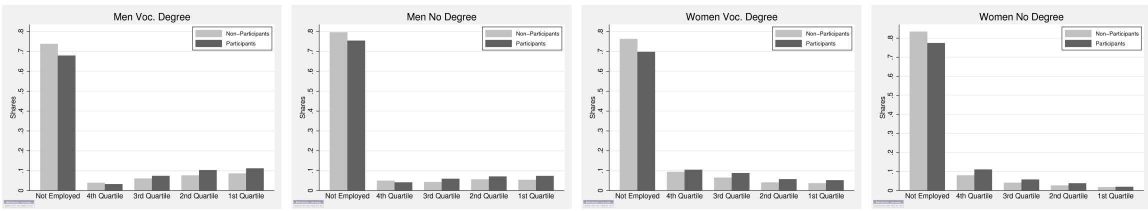
(a) Type 1




(b) Type 2




(c) Type 3




*Note:* Quartiles are based on the distribution of monthly earnings in the matched samples, aggregated across program types. 4th quartile: gross monthly earnings <1,050 Euros; 3rd quartile: gross monthly earnings 1,050–1,439 Euros; 2nd quartile: gross monthly earnings 1,440–1,890 Euros; 1st quartile: gross monthly earnings >1,890 Euros.

Figure 2.9c). Nevertheless, we can distinguish two clusters: for female participants with a vocational degree and male participants in general, additional jobs are generated in the top three quartiles of the monthly earnings distribution (with a tendency towards the first and second quartile for men with a degree, a slight tendency towards the top quartile for men without a degree, and a tendency towards the second and third quartile for women with a degree). Women without a vocational degree, however, find additional jobs in the three bottom quartiles, and especially in the fourth quartile. But this happens—and that is important—without a reduction of the share of individuals in the first quartile.

### 2.4.3 Treatment Effects: Age

For the analysis of the employment effects of training programs with respect to age, we distinguish three roughly equally sized age groups: individuals below 34 years, between 34 and 42 years, and above 42 years. Again, we first show effects on employment probabilities, and subsequently assess the impact of training programs on the monthly earnings distribution two years after program entry.

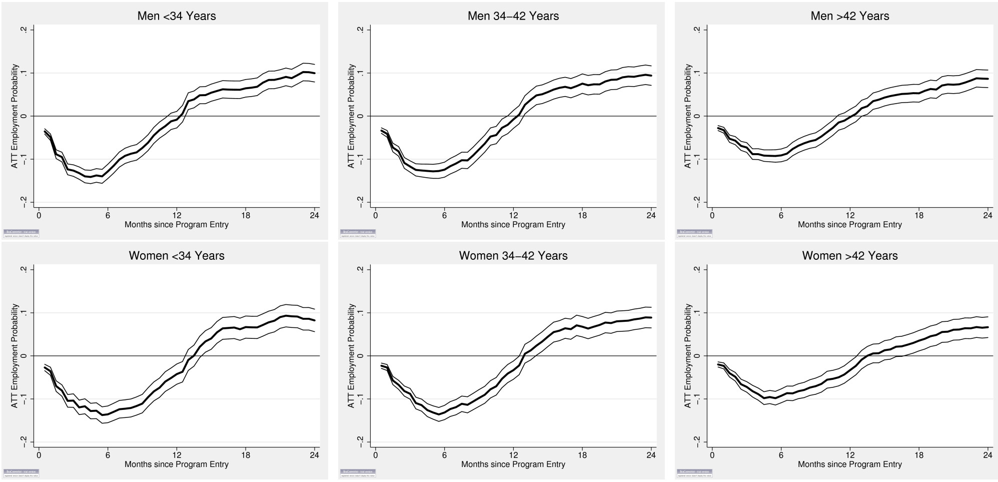
**Employment Probabilities** The general impression also carries over as far as the analysis of treatment effects with respect to age is concerned: the extent to which the ATT on employment probabilities vary between the sub-groups under consideration is quite small, and these differences are in almost all cases not significant.

More specifically, for type 1 the estimated treatment effects are very similar across the age groups under consideration (see Figure 2.10a and Table 2.6). Nonetheless, we calculate lower estimates for women in general, and especially for women who are at least 43 years old. For this sub-group, two years after program entry the point estimates are between 1.6 and 3.3 percentage points lower than the estimates for the other sub-groups. However, the ATT two years after program entry are in general not significantly different across sub-groups—and in particular not within the same gender. Only if we compare the point estimates for the youngest group of male and for the oldest group of female participants, we find a significant difference.

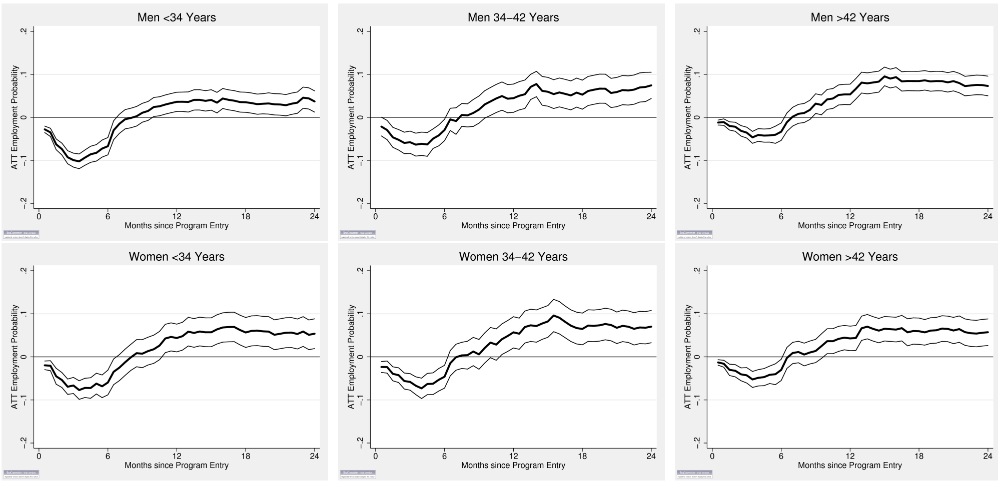
The treatment effects for type 2, likewise, exhibit in general no significant differ-

Figure 2.10: ATT employment probabilities by age group

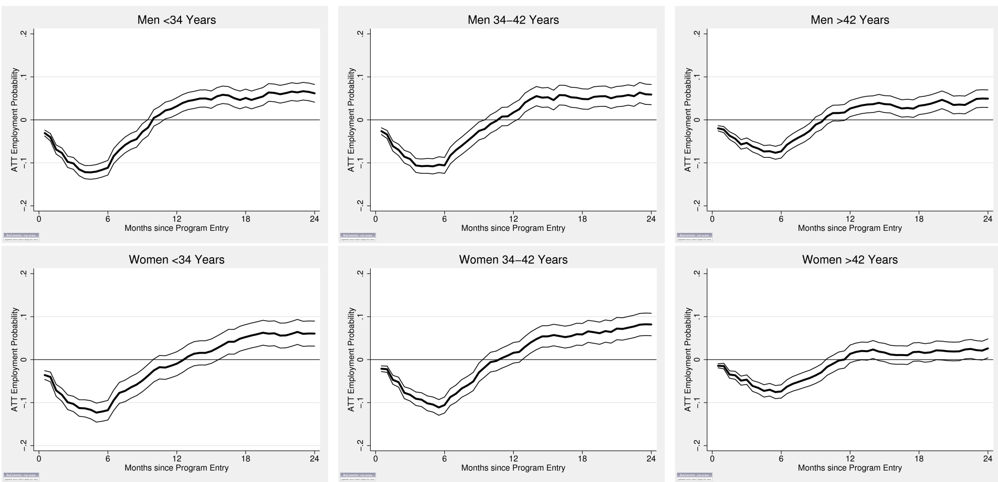
(a) Type 1




(b) Type 2




(c) Type 3




*Note:* Thick lines are point estimates of the ATT for the respective sub-group, while thin lines represent 95 percent confidence intervals.

Table 2.6: ATT employment probabilities by age group

| Type | Sex | Age Group | Month After Program Entry | Emp. Prob. NP | Emp. Prob. P | $\Delta_{ATT}$ | Lower Bound 95% CI | Upper Bound 95% CI |
|---|---|---|---|---|---|---|---|---|
| 1 | Female | <34 years | 6 | 0.2207 | 0.0850 | -0.1357 | -0.1551 | -0.1163 |
| | | | 12 | 0.2683 | 0.2260 | -0.0423 | -0.0658 | -0.0188 |
| | | | 18 | 0.2883 | 0.3551 | 0.0668 | 0.0411 | 0.0925 |
| | | | 24 | 0.3047 | 0.3871 | 0.0823 | 0.0562 | 0.1085 |
| 1 | Male | <34 years | 6 | 0.2605 | 0.1323 | -0.1283 | -0.1448 | -0.1117 |
| | | | 12 | 0.3033 | 0.2951 | -0.0082 | -0.0274 | 0.0109 |
| | | | 18 | 0.3524 | 0.4170 | 0.0646 | 0.0441 | 0.0850 |
| | | | 24 | 0.3394 | 0.4390 | 0.0997 | 0.0793 | 0.1200 |
| 1 | Female | 34–42 years | 6 | 0.1886 | 0.0571 | -0.1315 | -0.1480 | -0.1149 |
| | | | 12 | 0.2362 | 0.2031 | -0.0331 | -0.0543 | -0.0119 |
| | | | 18 | 0.2733 | 0.3410 | 0.0677 | 0.0441 | 0.0914 |
| | | | 24 | 0.2808 | 0.3697 | 0.0889 | 0.0650 | 0.1128 |
| 1 | Male | 34–42 years | 6 | 0.2243 | 0.0999 | -0.1244 | -0.1418 | -0.1069 |
| | | | 12 | 0.2607 | 0.2532 | -0.0075 | -0.0283 | 0.0133 |
| | | | 18 | 0.2985 | 0.3739 | 0.0754 | 0.0529 | 0.0979 |
| | | | 24 | 0.2966 | 0.3906 | 0.0940 | 0.0714 | 0.1167 |
| 1 | Female | >42 years | 6 | 0.1417 | 0.0490 | -0.0928 | -0.1089 | -0.0767 |
| | | | 12 | 0.1873 | 0.1543 | -0.0330 | -0.0536 | -0.0125 |
| | | | 18 | 0.2268 | 0.2621 | 0.0353 | 0.0116 | 0.0591 |
| | | | 24 | 0.2224 | 0.2890 | 0.0666 | 0.0426 | 0.0906 |
| 1 | Male | >42 years | 6 | 0.1624 | 0.0712 | -0.0912 | -0.1060 | -0.0763 |
| | | | 12 | 0.1892 | 0.1870 | -0.0022 | -0.0202 | 0.0158 |
| | | | 18 | 0.2364 | 0.2894 | 0.0530 | 0.0324 | 0.0736 |
| | | | 24 | 0.2184 | 0.3050 | 0.0866 | 0.0662 | 0.1070 |
| 2 | Female | <34 years | 6 | 0.2168 | 0.1573 | -0.0595 | -0.0882 | -0.0308 |
| | | | 12 | 0.2511 | 0.2949 | 0.0438 | 0.0116 | 0.0760 |
| | | | 18 | 0.2965 | 0.3532 | 0.0566 | 0.0219 | 0.0914 |
| | | | 24 | 0.2963 | 0.3501 | 0.0538 | 0.0192 | 0.0883 |
| 2 | Male | <34 years | 6 | 0.2505 | 0.1834 | -0.0671 | -0.0872 | -0.0469 |
| | | | 12 | 0.2791 | 0.3153 | 0.0362 | 0.0139 | 0.0585 |
| | | | 18 | 0.3238 | 0.3590 | 0.0352 | 0.0118 | 0.0586 |
| | | | 24 | 0.3313 | 0.3683 | 0.0370 | 0.0123 | 0.0617 |
| 2 | Female | 34–42 years | 6 | 0.1472 | 0.1012 | -0.0461 | -0.0725 | -0.0196 |
| | | | 12 | 0.2059 | 0.2628 | 0.0569 | 0.0206 | 0.0932 |
| | | | 18 | 0.2304 | 0.2953 | 0.0648 | 0.0280 | 0.1017 |
| | | | 24 | 0.2274 | 0.2976 | 0.0702 | 0.0328 | 0.1077 |
| 2 | Male | 34–42 years | 6 | 0.1541 | 0.1248 | -0.0293 | -0.0542 | -0.0044 |
| | | | 12 | 0.1905 | 0.2352 | 0.0446 | 0.0118 | 0.0775 |
| | | | 18 | 0.2186 | 0.2714 | 0.0528 | 0.0190 | 0.0866 |
| | | | 24 | 0.2109 | 0.2854 | 0.0745 | 0.0439 | 0.1051 |
| 2 | Female | >42 years | 6 | 0.1147 | 0.0845 | -0.0303 | -0.0549 | -0.0056 |
| | | | 12 | 0.1449 | 0.1878 | 0.0428 | 0.0143 | 0.0713 |
| | | | 18 | 0.1624 | 0.2219 | 0.0595 | 0.0290 | 0.0901 |
| | | | 24 | 0.1717 | 0.2289 | 0.0572 | 0.0263 | 0.0881 |
| 2 | Male | >42 years | 6 | 0.1161 | 0.0828 | -0.0333 | -0.0536 | -0.0130 |
| | | | 12 | 0.1342 | 0.1876 | 0.0534 | 0.0295 | 0.0772 |
| | | | 18 | 0.1498 | 0.2343 | 0.0845 | 0.0617 | 0.1072 |
| | | | 24 | 0.1516 | 0.2247 | 0.0731 | 0.0503 | 0.0959 |
| 3 | Female | <34 years | 6 | 0.2296 | 0.1119 | -0.1177 | -0.1405 | -0.0948 |
| | | | 12 | 0.2851 | 0.2756 | -0.0095 | -0.0372 | 0.0183 |
| | | | 18 | 0.2992 | 0.3518 | 0.0526 | 0.0237 | 0.0816 |
| | | | 24 | 0.3047 | 0.3654 | 0.0606 | 0.0315 | 0.0897 |
| 3 | Male | <34 years | 6 | 0.2475 | 0.1362 | -0.1114 | -0.1283 | -0.0944 |
| | | | 12 | 0.2830 | 0.3147 | 0.0318 | 0.0122 | 0.0513 |
| | | | 18 | 0.3303 | 0.3812 | 0.0509 | 0.0303 | 0.0715 |
| | | | 24 | 0.3246 | 0.3864 | 0.0618 | 0.0412 | 0.0823 |
| 3 | Female | 34–42 years | 6 | 0.1894 | 0.0838 | -0.1056 | -0.1242 | -0.0870 |
| | | | 12 | 0.2450 | 0.2609 | 0.0159 | -0.0080 | 0.0399 |
| | | | 18 | 0.2720 | 0.3306 | 0.0586 | 0.0331 | 0.0842 |
| | | | 24 | 0.2796 | 0.3614 | 0.0818 | 0.0557 | 0.1080 |
| 3 | Male | 34–42 years | 6 | 0.2150 | 0.1091 | -0.1059 | -0.1245 | -0.0874 |
| | | | 12 | 0.2387 | 0.2560 | 0.0173 | -0.0044 | 0.0389 |
| | | | 18 | 0.2902 | 0.3389 | 0.0487 | 0.0252 | 0.0722 |
| | | | 24 | 0.2865 | 0.3454 | 0.0589 | 0.0354 | 0.0824 |
| 3 | Female | >42 years | 6 | 0.1390 | 0.0650 | -0.0740 | -0.0894 | -0.0586 |
| | | | 12 | 0.1841 | 0.1975 | 0.0134 | -0.0071 | 0.0339 |
| | | | 18 | 0.2175 | 0.2358 | 0.0182 | -0.0032 | 0.0397 |
| | | | 24 | 0.2312 | 0.2574 | 0.0262 | 0.0042 | 0.0482 |
| 3 | Male | >42 years | 6 | 0.1595 | 0.0863 | -0.0733 | -0.0887 | -0.0578 |
| | | | 12 | 0.1860 | 0.2124 | 0.0264 | 0.0077 | 0.0452 |
| | | | 18 | 0.2295 | 0.2622 | 0.0328 | 0.0123 | 0.0533 |
| | | | 24 | 0.2227 | 0.2720 | 0.0494 | 0.0289 | 0.0698 |

*Note:* NP: Non-Participants; P: Participants; CI: confidence interval.

ences across sub-groups two years after program entry (see Figure 2.10b and Table 2.6). An exception applies for male participants, where we find a significantly lower point estimate for participants below 34 years if compared to those above 42 years. Moreover, while the overall picture suggests higher ATT for men than for women, an exception is the age group below 34 years. The ATT two years after program entry for men in this age group is particularly low (3.7 percentage points). But also if female participants in this age group are considered, treatment effects are relatively low.

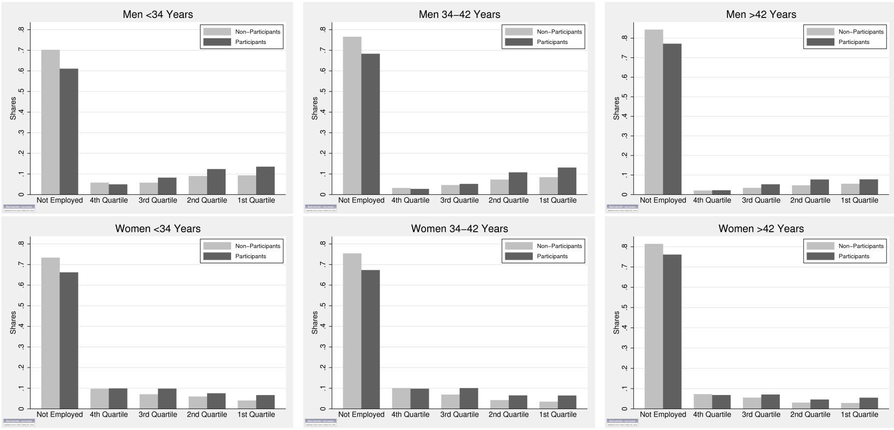
For type 3, we estimate relatively low—but still significantly positive—treatment effects for women above 42 years two years after program entry (see Figure 2.10c and Table 2.6). The estimated effects for male participants in the same age-group are also lower than in the other sub-groups, for which the ATT lie between 5.8 and 8.1 percentage points. The lock-in effects of program participation seem to be less persistent for men, as across all age groups the ATT become positive after around 9 months for men compared to 10–13 months for women. However, two years after program entry we calculate significantly different treatment effects compared to other sub-groups only for female participants above 42 years. The point estimate for this sub-group is significantly lower compared to men below 34 years and women between 34 and 42 years.

**Earnings Effects** The impact of participation in public training programs on the monthly earnings distribution two years after program entry is depicted in Figure 2.11. The overall picture suggests that the share of participants in the upper quartiles of the earnings distribution is generally higher than the share of matched non-participants, while this is for most sub-groups—and in particular as far as male individuals are considered—not the case in the bottom quartile.

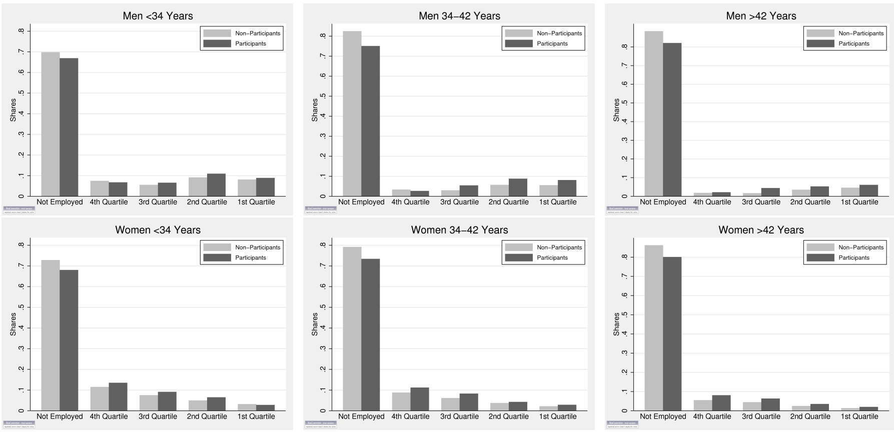
For type 1, the share of participants which is located in the top quartile of the monthly earnings distribution two years after program entry is across all sub-groups significantly higher than the respective share of comparison individuals (see Figure 2.11a). On the other hand, the differences between the shares in the bottom quartile of the earnings distribution are not significant. The shares of participants in the second and third quartile of the earnings distribution are across all sub-groups higher for participants than for matched non-participants.

Figure 2.11: Earnings distribution and employment effects 24 months after program entry by age group

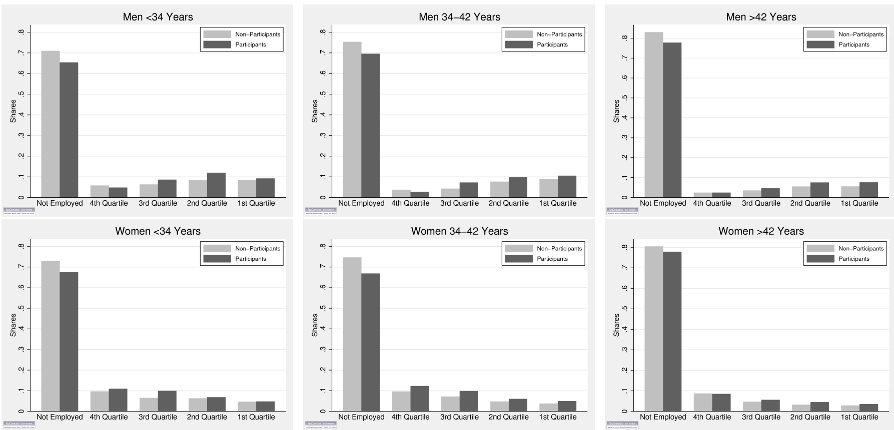
(a) Type 1




(b) Type 2




(c) Type 3




*Note:* Quartiles are based on the distribution of monthly earnings in the matched samples, aggregated across program types. 4th quartile: gross monthly earnings <1,050 Euros; 3rd quartile: gross monthly earnings 1,050–1,439 Euros; 2nd quartile: gross monthly earnings 1,440–1,920 Euros; 1st quartile: gross monthly earnings >1,920 Euros.

For types 2 and 3, the overall picture is less consistent than for type 1. Although the share of participants in the top quartile of the monthly earnings distribution is generally higher than the share of non-participants, we find significantly increased fractions only for male participants between 34 and 42 years as well as above 42 years (for both types). On the other hand, the share of male participants in type 2 between 34 and 42 years is significantly lower in the bottom quartile than the corresponding share of controls. Two other sub-groups exhibit a significantly higher share of treated individuals in the bottom quartile: female participants in type 2 above 42 years and male participants in type 3 between 34 and 42 years.

### 2.4.4 Sensitivity Analysis

To assess the sensitivity of our results with respect to the matching method, we additionally employ some alternative algorithms. Besides nearest neighbor matching without replacement, on which the above described results are based on, we calculate treatment effects based on (a) nearest neighbor matching with replacement, (b) caliper matching without replacement (with a maximum tolerance level of 0.001), and (c) radius matching (with a maximum tolerance level of 0.001). The results based on these three procedures reflect those presented above very closely. This is in line with Mueser et al. (2007) who also report quite similar results across a variety of matching methods if these methods are based on the same set of control variables.

As mentioned earlier, one could in principle choose a stricter definition of non-participation. Lechner and Wunsch (2008), for instance, distinguish participants from persons of the control group by conditioning on future non-participation. In their study, the impact of participation on employment probabilities two years after program entry is negative for most analyzed types. If we use a similar definition of non-participation, we find that this has an impact on the results presented here: depending on the respective sub-group, two years after program entry employment effects are 1.3–5.5 (mean: 3.8) percentage points lower for type 1, 0.1–4.6 (2.2) for type 2, and 0.0–6.0 (2.5) for type 3.

## 2.5 Conclusion

This chapter studies the effects of participation in public training programs for the unemployed in Germany. We apply propensity score matching methods and estimate the treatment effects for participants in the year 2002 using a rich administrative data set. We focus, next to average treatment effects on the treated, on treatment effects for different sub-groups of participants with respect to vocational education and age.

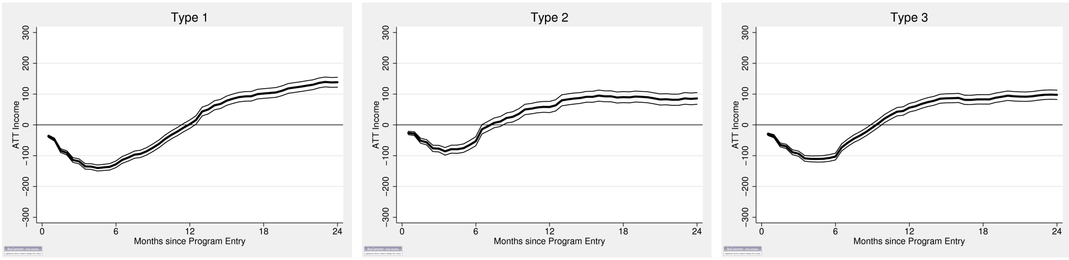
Considering three medium-term program types—with a median duration between 6 and 8 months and together accounting for roughly 85 percent of all participants in public training programs—our results indicate that program participation has a positive impact on employment probabilities for all sub-groups and program types. Moreover, participants seem to find more often higher paid jobs than non-participants. We present only little evidence for the presence of heterogeneous treatment effects, and the magnitude of these difference is quite small.

As far as the most important program type is concerned, we do not identify significant differences in treatment effects two years after entering the program across sub-groups of the same gender with respect to vocational eduction and age. Only if we compare sub-groups of male and female participants with each other, we find a significantly different ATT between the sub-groups of young men and old women. Also in case of this program type, the lock-in effect is remarkably shorter for male participants without a vocational degree. Similar results are found for the remaining two program types. Therefore, the overall picture suggests quite homogenous effects of program participation across sub-groups.

Our results are thus—at least in part—conflicting with the strategy to increasingly provide training to individuals with better employment prospects. This strategy has been implemented in Germany as a part of the reform of ALMP in 2003. After the reform, the caseworkers are asked to evaluate the employment prospects of the unemployed in advance and provide training only to individuals with a relatively high probability of entering employment after training participation. This does not take into account the relative gain compared to the situation without training. Although we find some evidence for a complementary relationship between advantageous employment prospects and the effectiveness of training in specific cases, our finding of positive treatment effects for all sub-groups raises the question whether the exclusion of 'bad' risks from training programs is a good strategy to reduce unemployment.

## 2.6 Appendix

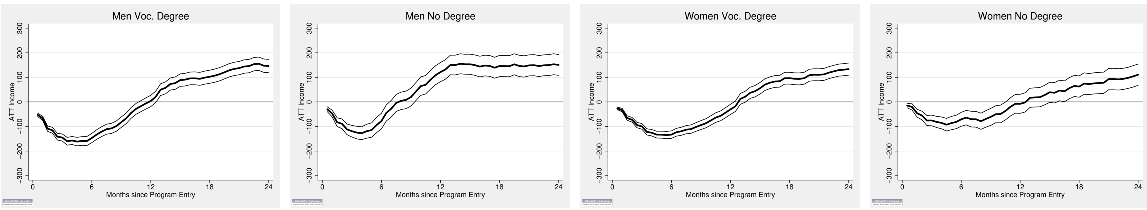
Figure 2.12: ATT monthly earnings



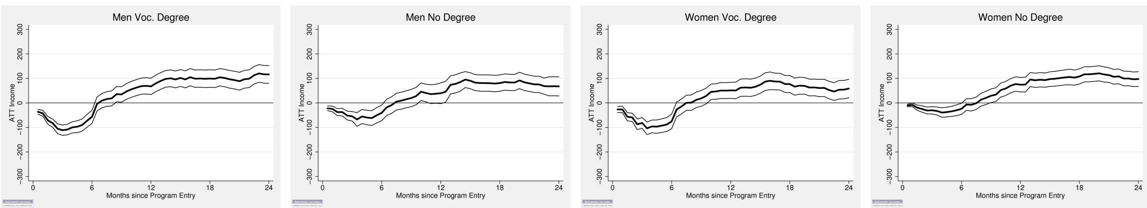

*Note:* Gross monthly earnings from employment (in Euros). Thick lines are point estimates of the ATT based on aggregated matched sub-samples with respect to vocational education, while thin lines represent 95 percent confidence intervals. The ATT for for the aggregated matched sub-samples with respect to age look very similar and are thus not displayed.

Figure 2.13: ATT monthly earnings by vocational education

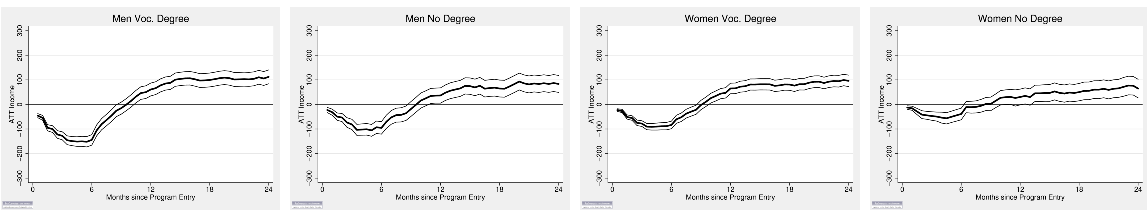
(a) Type 1




(b) Type 2



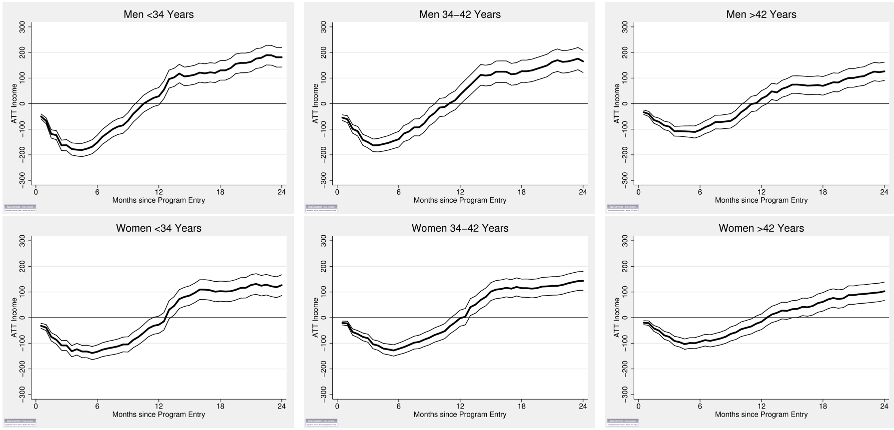

(c) Type 3



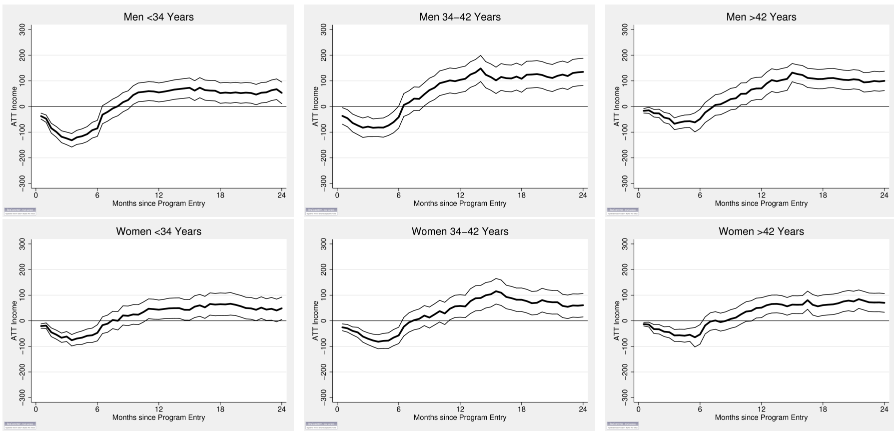

*Note:* Gross monthly earnings from employment (in Euros). Thick lines are point estimates of the ATT for the respective sub-group, while thin lines represent 95 percent confidence intervals.

Figure 2.14: ATT monthly earnings by age group

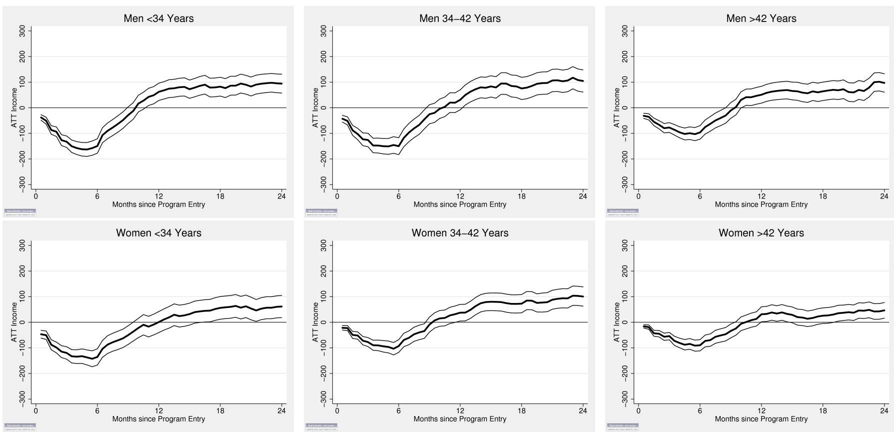
(a) Type 1




(b) Type 2




(c) Type 3




*Note:* Gross monthly earnings from employment (in Euros). Thick lines are point estimates of the ATT for the respective sub-group, while thin lines represent 95 percent confidence intervals.

Table 2.7: ATT monthly earnings by vocational education

| Type | Sex | Vocational Education | Month After Program Entry | Av. Earnings NP | Av. Earnings P | $\Delta_{ATT}$ | Lower Bound 95% CI | Upper Bound 95% CI |
|---|---|---|---|---|---|---|---|---|
| 1 | Female | Voc. Degree | 6 | 206.62 | 73.63 | -132.98 | -148.14 | -117.82 |
| | | | 12 | 271.15 | 240.99 | - 30.16 | - 50.94 | - 9.39 |
| | | | 18 | 307.64 | 404.26 | 96.62 | 72.70 | 120.54 |
| | | | 24 | 314.27 | 447.83 | 133.55 | 108.82 | 158.29 |
| 1 | Male | Voc. Degree | 6 | 301.29 | 154.32 | -146.97 | -165.89 | -128.06 |
| | | | 12 | 348.20 | 351.97 | 3.77 | - 19.18 | 26.73 |
| | | | 18 | 430.06 | 532.47 | 102.42 | 75.73 | 129.10 |
| | | | 24 | 428.86 | 574.88 | 146.01 | 118.88 | 173.14 |
| 1 | Female | No Degree | 6 | 140.04 | 69.45 | - 70.59 | - 98.94 | - 42.25 |
| | | | 12 | 184.26 | 176.59 | - 7.67 | - 46.36 | 31.03 |
| | | | 18 | 209.20 | 271.65 | 62.45 | 19.13 | 105.77 |
| | | | 24 | 182.62 | 293.35 | 110.73 | 67.56 | 153.90 |
| 1 | Male | No Degree | 6 | 220.74 | 142.92 | - 77.82 | -109.48 | - 46.15 |
| | | | 12 | 250.81 | 372.23 | 121.42 | 82.30 | 160.53 |
| | | | 18 | 309.02 | 455.23 | 146.21 | 103.00 | 189.42 |
| | | | 24 | 285.27 | 435.84 | 150.56 | 108.34 | 192.79 |
| 2 | Female | Voc. Degree | 6 | 217.05 | 141.94 | - 75.11 | -104.64 | - 45.59 |
| | | | 12 | 244.40 | 295.64 | 51.24 | 18.55 | 83.93 |
| | | | 18 | 274.34 | 351.05 | 76.72 | 42.00 | 111.43 |
| | | | 24 | 291.46 | 350.17 | 58.71 | 21.22 | 96.19 |
| 2 | Male | Voc. Degree | 6 | 265.78 | 208.89 | - 56.89 | - 83.52 | - 30.27 |
| | | | 12 | 343.30 | 410.83 | 67.53 | 34.05 | 101.01 |
| | | | 18 | 379.53 | 478.85 | 99.33 | 63.44 | 135.22 |
| | | | 24 | 368.83 | 485.06 | 116.23 | 80.14 | 152.33 |
| 2 | Female | No Degree | 6 | 113.04 | 89.00 | - 24.04 | - 46.70 | - 1.39 |
| | | | 12 | 138.12 | 212.54 | 74.42 | 44.93 | 103.92 |
| | | | 18 | 145.24 | 253.26 | 108.02 | 77.11 | 138.92 |
| | | | 24 | 143.60 | 240.91 | 97.31 | 67.16 | 127.46 |
| 2 | Male | No Degree | 6 | 186.05 | 145.52 | - 40.53 | - 72.28 | - 8.79 |
| | | | 12 | 236.68 | 276.13 | 39.45 | - 2.94 | 81.85 |
| | | | 18 | 239.31 | 320.16 | 80.85 | 47.64 | 114.07 |
| | | | 24 | 254.06 | 321.61 | 67.55 | 28.61 | 106.50 |
| 3 | Female | Voc. Degree | 6 | 186.90 | 103.45 | - 83.45 | - 98.86 | - 68.04 |
| | | | 12 | 235.43 | 300.43 | 65.00 | 43.54 | 86.46 |
| | | | 18 | 287.65 | 368.59 | 80.94 | 57.41 | 104.47 |
| | | | 24 | 295.52 | 391.46 | 95.94 | 72.63 | 119.24 |
| 3 | Male | Voc. Degree | 6 | 323.77 | 179.46 | -144.31 | -165.72 | -122.89 |
| | | | 12 | 364.77 | 425.51 | 60.74 | 34.67 | 86.81 |
| | | | 18 | 432.25 | 531.76 | 99.50 | 71.32 | 127.69 |
| | | | 24 | 426.26 | 538.42 | 112.16 | 83.98 | 140.34 |
| 3 | Female | No Degree | 6 | 107.69 | 68.91 | - 38.78 | - 62.45 | - 15.11 |
| | | | 12 | 158.26 | 188.52 | 30.27 | - 1.81 | 62.35 |
| | | | 18 | 180.54 | 230.66 | 50.12 | 14.85 | 85.38 |
| | | | 24 | 187.09 | 251.54 | 64.44 | 26.67 | 102.22 |
| 3 | Male | No Degree | 6 | 217.35 | 121.92 | - 95.43 | -120.73 | - 70.12 |
| | | | 12 | 259.94 | 296.25 | 36.31 | 5.26 | 67.36 |
| | | | 18 | 308.17 | 372.86 | 64.68 | 30.34 | 99.03 |
| | | | 24 | 295.70 | 379.05 | 83.35 | 48.86 | 117.84 |

*Note:* NP: Non-Participants; P: Participants; CI: confidence interval.

Table 2.8: ATT monthly earnings by age group

| Type | Sex | Age Group | Month After Program Entry | Av. Earnings NP | Av. Earnings P | $\Delta_{ATT}$ | Lower Bound 95% CI | Upper Bound 95% CI |
|---|---|---|---|---|---|---|---|---|
| 1 | Female | <34 years | 6 | 230.11 | 97.21 | -132.90 | -159.09 | -106.70 |
| | | | 12 | 290.27 | 262.67 | - 27.60 | - 61.13 | 5.93 |
| | | | 18 | 327.22 | 430.52 | 103.30 | 64.15 | 142.44 |
| | | | 24 | 341.62 | 468.31 | 126.68 | 86.22 | 167.14 |
| 1 | Male | <34 years | 6 | 348.33 | 196.91 | -151.42 | -179.65 | -123.20 |
| | | | 12 | 407.36 | 436.26 | 28.91 | - 5.18 | 62.99 |
| | | | 18 | 482.51 | 612.57 | 130.06 | 91.97 | 168.16 |
| | | | 24 | 469.87 | 651.28 | 181.41 | 143.15 | 219.67 |
| 1 | Female | 34–42 years | 6 | 192.55 | 70.86 | -121.68 | -144.20 | - 99.16 |
| | | | 12 | 250.53 | 248.48 | - 2.05 | - 33.48 | 29.38 |
| | | | 18 | 289.86 | 404.89 | 115.03 | 79.76 | 150.30 |
| | | | 24 | 305.28 | 448.63 | 143.36 | 106.68 | 180.03 |
| 1 | Male | 34–42 years | 6 | 286.37 | 147.17 | -139.19 | -169.65 | -108.73 |
| | | | 12 | 339.88 | 374.60 | 34.71 | - 3.04 | 72.47 |
| | | | 18 | 419.05 | 545.77 | 126.73 | 83.20 | 170.25 |
| | | | 24 | 398.83 | 563.95 | 165.12 | 121.73 | 208.52 |
| 1 | Female | >42 years | 6 | 148.89 | 54.85 | - 94.04 | -115.99 | - 72.09 |
| | | | 12 | 202.40 | 186.26 | - 16.14 | - 46.64 | 14.36 |
| | | | 18 | 248.02 | 309.62 | 61.60 | 25.66 | 97.54 |
| | | | 24 | 233.22 | 336.15 | 102.92 | 67.04 | 138.80 |
| 1 | Male | >42 years | 6 | 198.81 | 95.48 | -103.33 | -127.85 | - 78.82 |
| | | | 12 | 227.02 | 246.83 | 19.81 | - 10.98 | 50.60 |
| | | | 18 | 291.35 | 360.97 | 69.62 | 33.21 | 106.03 |
| | | | 24 | 261.82 | 387.98 | 126.16 | 90.33 | 161.98 |
| 2 | Female | <34 years | 6 | 214.08 | 166.92 | - 47.16 | - 78.61 | - 15.71 |
| | | | 12 | 263.81 | 307.16 | 43.35 | 5.80 | 80.91 |
| | | | 18 | 314.20 | 378.84 | 64.64 | 20.59 | 108.69 |
| | | | 24 | 322.87 | 371.54 | 48.66 | 4.96 | 92.36 |
| 2 | Male | <34 years | 6 | 323.49 | 238.24 | - 85.25 | -116.47 | - 54.04 |
| | | | 12 | 381.42 | 436.24 | 54.82 | 17.81 | 91.84 |
| | | | 18 | 451.24 | 503.54 | 52.30 | 12.58 | 92.02 |
| | | | 24 | 453.50 | 506.60 | 53.10 | 10.89 | 95.32 |
| 2 | Female | 34–42 years | 6 | 163.17 | 105.94 | - 57.23 | - 89.17 | - 25.29 |
| | | | 12 | 224.67 | 282.50 | 57.83 | 13.98 | 101.68 |
| | | | 18 | 249.86 | 331.91 | 82.05 | 36.04 | 128.06 |
| | | | 24 | 248.05 | 308.75 | 60.70 | 15.07 | 106.32 |
| 2 | Male | 34–42 years | 6 | 218.57 | 177.07 | - 41.51 | - 84.30 | 1.29 |
| | | | 12 | 246.42 | 350.36 | 103.95 | 54.98 | 152.92 |
| | | | 18 | 292.89 | 401.05 | 108.16 | 55.57 | 160.76 |
| | | | 24 | 279.34 | 414.18 | 134.84 | 81.67 | 188.01 |
| 2 | Female | >42 years | 6 | 133.32 | 80.11 | - 53.21 | - 91.31 | - 15.12 |
| | | | 12 | 154.05 | 205.20 | 51.15 | 12.15 | 90.15 |
| | | | 18 | 171.54 | 232.48 | 60.94 | 20.81 | 101.08 |
| | | | 24 | 165.93 | 235.94 | 70.01 | 33.32 | 106.70 |
| 2 | Male | >42 years | 6 | 133.32 | 80.11 | - 53.21 | - 91.31 | - 15.12 |
| | | | 12 | 154.05 | 205.20 | 51.15 | 12.15 | 90.15 |
| | | | 18 | 171.54 | 232.48 | 60.94 | 20.81 | 101.08 |
| | | | 24 | 165.93 | 235.94 | 70.01 | 33.32 | 106.70 |
| 3 | Female | <34 years | 6 | 260.80 | 125.36 | -135.43 | -167.16 | -103.71 |
| | | | 12 | 322.88 | 323.65 | 0.77 | - 40.28 | 41.82 |
| | | | 18 | 359.07 | 415.25 | 56.18 | 11.88 | 100.47 |
| | | | 24 | 354.53 | 415.92 | 61.40 | 18.05 | 104.74 |
| 3 | Male | <34 years | 6 | 339.97 | 190.56 | -149.41 | -177.57 | -121.24 |
| | | | 12 | 387.92 | 450.29 | 62.37 | 28.58 | 96.17 |
| | | | 18 | 459.78 | 542.12 | 82.34 | 45.65 | 119.03 |
| | | | 24 | 452.25 | 546.53 | 94.28 | 57.50 | 131.06 |
| 3 | Female | 34–42 years | 6 | 199.68 | 106.70 | - 92.98 | -117.77 | - 68.19 |
| | | | 12 | 264.48 | 301.88 | 37.40 | 4.98 | 69.81 |
| | | | 18 | 303.22 | 375.79 | 72.57 | 36.93 | 108.21 |
| | | | 24 | 318.04 | 418.51 | 100.48 | 62.84 | 138.11 |
| 3 | Male | 34–42 years | 6 | 311.08 | 161.07 | -150.01 | -182.99 | -117.03 |
| | | | 12 | 348.39 | 379.27 | 30.88 | - 8.15 | 69.91 |
| | | | 18 | 435.88 | 511.31 | 75.43 | 32.06 | 118.81 |
| | | | 24 | 411.26 | 515.90 | 104.64 | 61.44 | 147.84 |
| 3 | Female | >42 years | 6 | 158.93 | 69.03 | - 89.90 | -111.77 | - 68.03 |
| | | | 12 | 196.95 | 228.15 | 31.21 | 1.99 | 60.41 |
| | | | 18 | 235.84 | 261.54 | 25.70 | - 4.75 | 56.14 |
| | | | 24 | 235.96 | 282.39 | 46.43 | 15.89 | 76.97 |
| 3 | Male | >42 years | 6 | 213.54 | 117.16 | - 96.38 | -122.77 | - 69.99 |
| | | | 12 | 246.03 | 296.89 | 50.86 | 17.97 | 83.75 |
| | | | 18 | 301.82 | 366.04 | 64.23 | 28.21 | 100.24 |
| | | | 24 | 278.16 | 375.06 | 96.90 | 61.09 | 132.70 |

*Note:* NP: Non-Participants; P: Participants; CI: confidence interval.

Table 2.9: Overall matching quality: Type 1

| | | *Vocational Education* | | | *Age* | | |
|---|---|---|---|---|---|---|---|
| | | Participants | Non-Participants | %-Bias | Participants | Non-Participants | %-Bias |
| Age | Before Matching | 37.142 | 39.389 | -21.0 | 37.142 | 39.389 | -21.0 |
| | After Matching | 37.151 | 37.051 | 0.9 | 37.158 | 37.138 | 0.2 |
| German | Before Matching | 0.0141 | 0.0254 | - 8.1 | 0.0141 | 0.0254 | - 8.1 |
| | After Matching | 0.0141 | 0.0146 | - 0.3 | 0.0141 | 0.0142 | - 0.1 |
| Married | Before Matching | 0.4780 | 0.4867 | - 1.7 | 0.4780 | 0.4867 | - 1.7 |
| | After Matching | 0.4783 | 0.4744 | 0.8 | 0.4785 | 0.4758 | 0.5 |
| Dependent children: youngest kid 0–3 years old | Before Matching | 0.0655 | 0.0496 | 6.8 | 0.0655 | 0.0496 | 6.8 |
| | After Matching | 0.0655 | 0.0650 | 0.2 | 0.0650 | 0.0653 | - 0.2 |
| Dependent children: youngest kid 4–14 years old | Before Matching | 0.2407 | 0.1915 | 12.0 | 0.2407 | 0.1915 | 12.0 |
| | After Matching | 0.2411 | 0.2436 | - 0.6 | 0.2411 | 0.2436 | - 0.6 |
| No graduation | Before Matching | 0.0569 | 0.1234 | -23.4 | 0.0569 | 0.1234 | -23.4 |
| | After Matching | 0.0568 | 0.0571 | - 0.1 | 0.0568 | 0.0579 | - 0.4 |
| First stage of secondary level | Before Matching | 0.3190 | 0.4747 | -32.2 | 0.3190 | 0.4747 | -32.2 |
| | After Matching | 0.3189 | 0.3175 | 0.3 | 0.3190 | 0.3137 | 1.1 |
| Second stage of secondary level | Before Matching | 0.4096 | 0.2833 | 26.8 | 0.4096 | 0.2833 | 26.8 |
| | After Matching | 0.4098 | 0.4075 | 0.5 | 0.4096 | 0.4105 | - 0.2 |
| Advanced technical college entrance qualification | Before Matching | 0.0682 | 0.0384 | 13.3 | 0.0682 | 0.0384 | 13.3 |
| | After Matching | 0.0681 | 0.0688 | - 0.3 | 0.0682 | 0.0700 | - 0.8 |
| General qualification for university entrance | Before Matching | 0.1463 | 0.0802 | 21.0 | 0.1463 | 0.0802 | 21.0 |
| | After Matching | 0.1464 | 0.1491 | - 0.8 | 0.1465 | 0.1479 | - 0.5 |
| No vocational degree | Before Matching | 0.1918 | 0.3651 | -39.4 | 0.1918 | 0.3651 | -39.4 |
| | After Matching | 0.1913 | 0.1913 | 0.0 | 0.1913 | 0.1924 | - 0.2 |
| In-plant training | Before Matching | 0.5946 | 0.5145 | 16.2 | 0.5946 | 0.5145 | 16.2 |
| | After Matching | 0.5953 | 0.5931 | 0.4 | 0.5954 | 0.5910 | 0.9 |
| Off-the-job training, vocational school, technical school | Before Matching | 0.1007 | 0.0693 | 11.3 | 0.1007 | 0.0693 | 11.3 |
| | After Matching | 0.1005 | 0.1010 | - 0.2 | 0.1006 | 0.1021 | - 0.6 |
| University, advanced technical college | Before Matching | 0.1129 | 0.0511 | 22.7 | 0.1129 | 0.0511 | 22.7 |
| | After Matching | 0.1129 | 0.1146 | - 0.7 | 0.1127 | 0.1145 | - 0.7 |
| Share of unemployment in 1st year before program entry | Before Matching | 0.5840 | 0.5989 | - 4.7 | 0.5840 | 0.5989 | - 4.7 |
| | After Matching | 0.5842 | 0.5738 | 3.3 | 0.5841 | 0.5757 | 2.6 |
| Share of unemployment in 2nd year before program entry | Before Matching | 0.2408 | 0.3252 | -24.2 | 0.2408 | 0.3252 | -24.2 |
| | After Matching | 0.2409 | 0.2386 | 0.7 | 0.2408 | 0.2399 | 0.3 |
| Share of unemployment in 3rd year before program entry | Before Matching | 0.1981 | 0.2875 | -25.9 | 0.1981 | 0.2875 | -25.9 |
| | After Matching | 0.1978 | 0.1968 | 0.3 | 0.1979 | 0.1992 | - 0.4 |
| Share of unemployment in 4th year before program entry | Before Matching | 0.1765 | 0.2575 | -24.3 | 0.1765 | 0.2575 | -24.3 |
| | After Matching | 0.1758 | 0.1760 | - 0.0 | 0.1758 | 0.1785 | - 0.8 |
| Share of employment in 1st year before program entry | Before Matching | 0.2425 | 0.2027 | 13.4 | 0.2425 | 0.2027 | 13.4 |
| | After Matching | 0.2426 | 0.2448 | - 0.7 | 0.2426 | 0.2440 | - 0.5 |
| Share of employment in 2nd year before program entry | Before Matching | 0.4681 | 0.3667 | 23.9 | 0.4681 | 0.3667 | 23.9 |
| | After Matching | 0.4687 | 0.4617 | 1.6 | 0.4685 | 0.4636 | 1.2 |
| Share of employment in 3rd year before program entry | Before Matching | 0.4781 | 0.3882 | 20.5 | 0.4781 | 0.3882 | 20.5 |
| | After Matching | 0.4791 | 0.4686 | 2.4 | 0.4791 | 0.4741 | 1.1 |
| Share of employment in 4th year before program entry | Before Matching | 0.4759 | 0.3994 | 17.4 | 0.4759 | 0.3994 | 17.4 |
| | After Matching | 0.4774 | 0.4709 | 1.5 | 0.4775 | 0.4733 | 0.9 |
| Pseudo-$R^2$ | Before Matching | | 0.058 | | | 0.058 | |
| | After Matching | | 0.001 | | | 0.000 | |
| Mean standardized bias | Before Matching | | 13.367 | | | 13.367 | |
| | After Matching | | 0.687 | | | 0.493 | |

*Note:* Statistics are based on aggregated sub-samples before and after matching (previously stratified with respect to vocational education and age). Only selected variables are reported, while the specifications include more variables. However, mean standardized bias and pseudo-$R^2$ refer to 75 variables at least included in the different specifications.

Table 2.10: Overall matching quality: Type 2

| | | *Vocational Education* | | | *Age* | | |
|---|---|---|---|---|---|---|---|
| | | Participants | Non-Participants | %-Bias | Participants | Non-Participants | %-Bias |
| Age | Before Matching | 35.698 | 39.389 | -32.3 | 35.698 | 39.389 | -32.3 |
| | After Matching | 35.673 | 35.822 | - 1.3 | 35.671 | 35.591 | 0.7 |
| German | Before Matching | 0.0164 | 0.0254 | - 6.3 | 0.0164 | 0.0254 | - 6.3 |
| | After Matching | 0.0165 | 0.0174 | - 0.6 | 0.0165 | 0.0169 | - 0.3 |
| Married | Before Matching | 0.4112 | 0.4867 | -15.2 | 0.4112 | 0.4867 | -15.2 |
| | After Matching | 0.4113 | 0.4137 | - 0.5 | 0.4112 | 0.4118 | - 0.1 |
| Dependent children: youngest kid 0–3 years old | Before Matching | 0.0486 | 0.0496 | - 0.5 | 0.0486 | 0.0496 | - 0.5 |
| | After Matching | 0.0485 | 0.0490 | - 0.2 | 0.0485 | 0.0450 | 1.6 |
| Dependent children: youngest kid 4–14 years old | Before Matching | 0.2082 | 0.1915 | 4.2 | 0.2082 | 0.1915 | 4.2 |
| | After Matching | 0.2087 | 0.2063 | 0.6 | 0.2087 | 0.2082 | 0.1 |
| No graduation | Before Matching | 0.1241 | 0.1234 | 0.2 | 0.1241 | 0.1235 | 0.2 |
| | After Matching | 0.1242 | 0.1255 | - 0.4 | 0.1243 | 0.1267 | - 0.7 |
| First stage of secondary level | Before Matching | 0.5325 | 0.4747 | 11.6 | 0.5325 | 0.4747 | 11.6 |
| | After Matching | 0.5324 | 0.5327 | - 0.1 | 0.5324 | 0.5298 | 0.5 |
| Second stage of secondary level | Before Matching | 0.2576 | 0.2833 | - 5.8 | 0.2576 | 0.2833 | - 5.8 |
| | After Matching | 0.2575 | 0.2534 | 0.9 | 0.2575 | 0.2558 | 0.4 |
| Advanced technical college entrance qualification | Before Matching | 0.0337 | 0.0384 | - 2.5 | 0.0337 | 0.0384 | - 2.5 |
| | After Matching | 0.0338 | 0.0360 | - 1.2 | 0.0338 | 0.0352 | - 0.7 |
| General qualification for university entrance | Before Matching | 0.0520 | 0.0801 | -11.3 | 0.0520 | 0.0801 | -11.3 |
| | After Matching | 0.0521 | 0.0524 | - 0.1 | 0.0521 | 0.0526 | - 0.2 |
| No vocational degree | Before Matching | 0.3862 | 0.3652 | 4.3 | 0.3862 | 0.3652 | 4.3 |
| | After Matching | 0.3856 | 0.3856 | 0.0 | 0.3855 | 0.3895 | - 0.8 |
| In-plant training | Before Matching | 0.5069 | 0.5145 | - 1.5 | 0.5069 | 0.5145 | - 1.5 |
| | After Matching | 0.5075 | 0.5074 | 0.0 | 0.5076 | 0.5024 | 1.0 |
| Off-the-job training, vocational school, technical school | Before Matching | 0.0766 | 0.0693 | 2.8 | 0.0766 | 0.0693 | 2.8 |
| | After Matching | 0.0766 | 0.0767 | - 0.0 | 0.0766 | 0.0767 | - 0.0 |
| University, advanced technical college | Before Matching | 0.0303 | 0.0511 | -10.6 | 0.0303 | 0.0511 | -10.6 |
| | After Matching | 0.0303 | 0.0303 | 0.0 | 0.0303 | 0.0315 | - 0.6 |
| Share of unemployment in 1st year before program entry | Before Matching | 0.6352 | 0.5989 | 11.4 | 0.6352 | 0.5989 | 11.4 |
| | After Matching | 0.6344 | 0.6285 | 1.9 | 0.6344 | 0.6274 | 2.2 |
| Share of unemployment in 2nd year before program entry | Before Matching | 0.3267 | 0.3252 | 0.4 | 0.3267 | 0.3252 | 0.4 |
| | After Matching | 0.3249 | 0.3259 | - 0.3 | 0.3249 | 0.3243 | 0.2 |
| Share of unemployment in 3rd year before program entry | Before Matching | 0.2688 | 0.2875 | - 5.0 | 0.2688 | 0.2875 | - 5.0 |
| | After Matching | 0.2666 | 0.2684 | - 0.5 | 0.2666 | 0.2682 | - 0.4 |
| Share of unemployment in 4th year before program entry | Before Matching | 0.2373 | 0.2575 | - 5.6 | 0.2373 | 0.2575 | - 5.6 |
| | After Matching | 0.2350 | 0.2347 | 0.1 | 0.2349 | 0.2368 | - 0.5 |
| Share of employment in 1st year before program entry | Before Matching | 0.1842 | 0.2027 | - 6.5 | 0.1842 | 0.2027 | - 6.5 |
| | After Matching | 0.1847 | 0.1833 | 0.5 | 0.1847 | 0.1845 | 0.1 |
| Share of employment in 2nd year before program entry | Before Matching | 0.3799 | 0.3666 | 3.2 | 0.3799 | 0.3666 | 3.2 |
| | After Matching | 0.3811 | 0.3745 | 1.6 | 0.3811 | 0.3771 | 1.0 |
| Share of employment in 3rd year before program entry | Before Matching | 0.3958 | 0.3882 | 1.8 | 0.3958 | 0.3881 | 1.8 |
| | After Matching | 0.3975 | 0.3935 | 0.9 | 0.3974 | 0.3940 | 0.8 |
| Share of employment in 4th year before program entry | Before Matching | 0.3818 | 0.3994 | - 4.1 | 0.3818 | 0.3994 | - 4.1 |
| | After Matching | 0.3834 | 0.3843 | - 0.2 | 0.3833 | 0.3814 | 0.5 |
| Pseudo-$R^2$ | Before Matching | | 0.076 | | | 0.076 | |
| | After Matching | | 0.001 | | | 0.001 | |
| Mean standardized bias | Before Matching | | 9.432 | | | 9.432 | |
| | After Matching | | 0.638 | | | 0.856 | |

*Note:* Statistics are based on aggregated sub-samples before and after matching (previously stratified with respect to vocational education and age). Only selected variables are reported, while the specifications include more variables. However, mean standardized bias and pseudo-$R^2$ refer to 75 variables at least included in the different specifications.

Table 2.11: Overall matching quality: Type 3

| | | *Vocational Education* | | | *Age* | | |
|---|---|---|---|---|---|---|---|
| | | Participants | Non-Participants | %-Bias | Participants | Non-Participants | %-Bias |
| Age | Before Matching | 37.782 | 39.389 | -14.8 | 37.782 | 39.389 | -14.8 |
| | After Matching | 37.786 | 37.802 | - 0.1 | 37.787 | 37.783 | 0.0 |
| German | Before Matching | 0.0149 | 0.0254 | - 7.5 | 0.0149 | 0.0254 | - 7.5 |
| | After Matching | 0.0149 | 0.0155 | - 0.4 | 0.0149 | 0.0163 | - 1.0 |
| Married | Before Matching | 0.4971 | 0.4867 | 2.1 | 0.4971 | 0.4867 | 2.1 |
| | After Matching | 0.4977 | 0.4942 | 0.7 | 0.4976 | 0.4967 | 0.2 |
| Dependent children: youngest kid 0–3 years old | Before Matching | 0.0601 | 0.0496 | 4.6 | 0.0601 | 0.0496 | 4.6 |
| | After Matching | 0.0601 | 0.0562 | 1.7 | 0.0601 | 0.0588 | 0.6 |
| Dependent children: youngest kid 4–14 years old | Before Matching | 0.2308 | 0.1915 | 9.7 | 0.2308 | 0.1915 | 9.7 |
| | After Matching | 0.2310 | 0.2298 | 0.3 | 0.2310 | 0.2329 | - 0.5 |
| No graduation | Before Matching | 0.0967 | 0.1234 | - 8.6 | 0.0967 | 0.1234 | - 8.6 |
| | After Matching | 0.0967 | 0.0933 | 1.1 | 0.0967 | 0.0949 | 0.6 |
| First stage of secondary level | Before Matching | 0.4290 | 0.4747 | - 9.2 | 0.4290 | 0.4747 | - 9.2 |
| | After Matching | 0.4288 | 0.4335 | - 0.9 | 0.4288 | 0.4326 | - 0.8 |
| Second stage of secondary level | Before Matching | 0.3890 | 0.2833 | 22.5 | 0.3890 | 0.2833 | 22.5 |
| | After Matching | 0.3892 | 0.3880 | 0.2 | 0.3892 | 0.3883 | 0.2 |
| Advanced technical college entrance qualification | Before Matching | 0.0322 | 0.0384 | - 3.4 | 0.0322 | 0.0384 | - 3.4 |
| | After Matching | 0.0323 | 0.0302 | 1.2 | 0.0323 | 0.0315 | 0.4 |
| General qualification for university entrance | Before Matching | 0.0531 | 0.0801 | -10.9 | 0.0531 | 0.0801 | -10.9 |
| | After Matching | 0.0531 | 0.0551 | - 0.8 | 0.0531 | 0.0528 | 0.1 |
| No vocational degree | Before Matching | 0.2722 | 0.3652 | -20.1 | 0.2722 | 0.3652 | -20.1 |
| | After Matching | 0.2715 | 0.2715 | - 0.0 | 0.2714 | 0.2720 | - 0.1 |
| In-plant training | Before Matching | 0.6170 | 0.5145 | 20.8 | 0.6170 | 0.5145 | 20.8 |
| | After Matching | 0.6177 | 0.6167 | 0.2 | 0.6177 | 0.6182 | - 0.1 |
| Off-the-job training, vocational school, technical school | Before Matching | 0.0792 | 0.0693 | 3.8 | 0.0792 | 0.0693 | 3.8 |
| | After Matching | 0.0793 | 0.0809 | - 0.6 | 0.0793 | 0.0792 | 0.1 |
| University, advanced technical college | Before Matching | 0.0316 | 0.0511 | - 9.8 | 0.0316 | 0.0511 | - 9.8 |
| | After Matching | 0.0316 | 0.0309 | 0.4 | 0.0316 | 0.0307 | 0.5 |
| Share of unemployment in 1st year before program entry | Before Matching | 0.5917 | 0.5989 | - 2.3 | 0.5917 | 0.5989 | - 2.3 |
| | After Matching | 0.5913 | 0.5858 | 1.7 | 0.5913 | 0.5835 | 2.5 |
| Share of unemployment in 2nd year before program entry | Before Matching | 0.2782 | 0.3252 | -13.0 | 0.2782 | 0.3252 | -13.0 |
| | After Matching | 0.2775 | 0.2771 | 0.1 | 0.2776 | 0.2764 | 0.3 |
| Share of unemployment in 3rd year before program entry | Before Matching | 0.2353 | 0.2875 | -14.6 | 0.2353 | 0.2875 | -14.6 |
| | After Matching | 0.2345 | 0.2354 | - 0.2 | 0.2345 | 0.2381 | - 1.0 |
| Share of unemployment in 4th year before program entry | Before Matching | 0.2128 | 0.2575 | -13.0 | 0.2128 | 0.2575 | -13.0 |
| | After Matching | 0.2118 | 0.2147 | - 0.8 | 0.2118 | 0.2159 | - 1.2 |
| Share of employment in 1st year before program entry | Before Matching | 0.2399 | 0.2027 | 12.6 | 0.2399 | 0.2027 | 12.6 |
| | After Matching | 0.2401 | 0.2434 | - 1.1 | 0.2400 | 0.2424 | - 0.8 |
| Share of employment in 2nd year before program entry | Before Matching | 0.4560 | 0.3667 | 21.0 | 0.4560 | 0.3666 | 21.0 |
| | After Matching | 0.4565 | 0.4559 | 0.1 | 0.4565 | 0.4531 | 0.8 |
| Share of employment in 3rd year before program entry | Before Matching | 0.4655 | 0.3882 | 17.7 | 0.4655 | 0.3881 | 17.7 |
| | After Matching | 0.4662 | 0.4632 | 0.7 | 0.4662 | 0.4605 | 1.3 |
| Share of employment in 4th year before program entry | Before Matching | 0.4630 | 0.3994 | 14.5 | 0.4630 | 0.3994 | 14.5 |
| | After Matching | 0.4640 | 0.4621 | 0.4 | 0.4641 | 0.4590 | 1.2 |
| Pseudo-$R^2$ | Before Matching | | 0.029 | | | 0.029 | |
| | After Matching | | 0.000 | | | 0.001 | |
| Mean standardized bias | Before Matching | | 9.904 | | | 9.905 | |
| | After Matching | | 0.582 | | | 0.615 | |

*Note:* Statistics are based on aggregated sub-samples before and after matching (previously stratified with respect to vocational education and age). Only selected variables are reported, while the specifications include more variables. However, mean standardized bias and pseudo-$R^2$ refer to 75 variables at least included in the different specifications.

# Chapter 3

# Vouchers and Caseworkers in Public Training Programs: Evidence from the Hartz Reform

*Vouchers are a common instrument in many fields of public services—in particular in the field of education. They are thus quite extensively studied in the literature, but the approach is novel in the context of delivering active labor market policy. This chapter analyzes the impacts of the introduction of vouchers for public training programs in Germany.*[23]

[23] This chapter is based on joint work with Arne Uhlendorff and Zhong Zhao (Rinne et al., 2008).

## 3.1 Introduction

Germany reformed its active labor market policy (ALMP) in a series of reforms which are commonly known as the Hartz reforms. When the first wave of these reforms (Hartz I/II) came into force on January 1, 2003 the provision of public training programs was substantially changed. The most important change was the introduction of a voucher scheme. The former contracting-out system was abandoned and replaced by a system in which job seekers are free to select their training provider in the market. Previously this choice was made by the caseworker. However, the participants are not completely free in their choice, because the content of the training is still assigned by the caseworker. Vouchers are a common instrument in other fields of public services, but this approach is rather novel in the context of delivering ALMP. In addition to the voucher scheme, a stricter selection of the participants by the caseworkers was introduced.

This reform provides us a valuable opportunity to study not only the overall training impact, which is the main parameter of interest in most of the literature on training, but also to investigate the effect of different components inside the training "black box". In this chapter, we estimate the impact of the reform on the effectiveness of the most important training program type for the unemployed in Germany. We disentangle the effect induced by the introduction of vouchers (voucher effect) from the effect induced by a more positive selection on the basis of expected employability by the caseworkers (selection effect).

Increased consumer choice and provider competition are the main arguments in favor of the introduction of vouchers (see, e.g., Steuerle, 2000). It is argued that allowing participants to choose the training provider in accordance to their preferences will lead to better matches between the unemployed and training providers, which will increase the effectiveness of participation. In addition, greater freedom of choice may encourage more competition among the providers. Training providers may have to compete more if they must regularly face the demand of participants instead of having a longer-term contract with the employment agency. This could lead to a further increase of the match quality and therewith of the effectiveness of training programs. On the other hand, there may be obstacles in case of public training programs which counteract the potential positive impacts of vouchers. Generally, it is argued that the consumer—in our case the unemployed—may lack competence or resources to optimally choose, and that information asymmetries may lead to choices which do not truly reflect consumers' preferences. For example, the caseworkers may know more about the availability of training courses and the quality of training providers than the unemployed, because of their experience with previous participants. For a discussion of advantages and disadvantages of training vouchers see Barnow (2000, 2009).

Vouchers have been widely used in other fields of public services—in particular in the field of education—and are quite extensively studied in the literature.[24] There exist some studies on vouchers for pre-school education (e.g., Viitanen, 2007), but most studies focus on school education (e.g., Manski, 1992; Nechyba, 2000; Angrist et al., 2002; Krueger and Zhu, 2004). Ladd (2002) presents a review of major studies on school vouchers. She concludes that the overall picture that can be drawn from these studies is rather inconclusive, and that the results are not very robust. For instance, studies with U.S. data typically indicate that there is insufficient information to draw clear conclusions about the net effects on 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. 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 exist only few studies of vouchers for job training programs. Barnow (2009) gives an overview of studies on vouchers in U.S. vocational training programs. The empirical evidence for the effectiveness of training vouchers for dislocated workers is mixed. To the best of our knowledge there exist no econometric study evaluating the effectiveness of the introduction of vouchers for training programs targeted at unemployed

[24]The discussion about vouchers in the educational context started with Friedman (1962).

individuals.[25]

In this chapter we focus on a voucher for training programs in the context of ALMP, and estimate the impact of the introduction of training vouchers on the effectiveness of the most important program type in Germany: occupation-related or general training. Participants in this program either learn specific skills required for a certain vocation or receive qualifications which are of general vocational use. In the pre-reform period about 60 percent of all unemployed participants in public training programs were assigned to this particular type, and it became even more important after the reform in 2003 as this share increased to more than 70 percent.

In addition to the training voucher, a stricter selection of participants and program types by the caseworkers based on the expected reemployment probability was introduced. This implies that the caseworkers tend to select individuals with higher reemployment probabilities for participation—independent of the individual gain resulting from participation. Thus, the overall effect of the reform could result from the introduction of the voucher system (voucher effect) and from a change in the composition of participants (selection effect). To decompose the overall reform effect into these two effects, we apply a two-step propensity score matching procedure using a rich administrative data set. This approach allows a comparison between pre- and post-reform participants who have similar observable characteristics. Furthermore, we apply regression analyses to the matched data to adjust for possible remaining unbalanced covariates, and to account for changes in the general economic conditions.

We find a slightly positive impact of the reform. The decomposition of this overall effect shows that the selection effect is—if at all—slightly negative. This finding is consistent with Lechner and Smith (2007). Although the selection effect in our chapter differs from the effect of caseworkers in their paper, the major part of our selection effect is also due to the caseworker. Our findings imply that using post-training outcomes as a performance standard is not a good strategy to improve the effectiveness and efficiency of public training programs; see Heckman et al. (1997) for

[25] 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. Winterhager et al. (2006) evaluate the effectiveness of this instrument and find a positive impact on the employment probability of voucher recipients in West Germany.

an insightful discussion on performance standards in the context of the Job Training Partnership Act. Furthermore, we find evidence that the voucher effect increased both the employment probability and earnings of the participants. This effect becomes substantially positive after around 6 months of training, and decreases slightly at the end of our observation period (1.5 years after program entry). Our results are mainly driven by skilled participants. We do not find any significant reform effect for the unskilled. While the former group can take advantage from an increased consumer sovereignty, unskilled individuals seem to have problems in adequately using the newly introduced voucher.

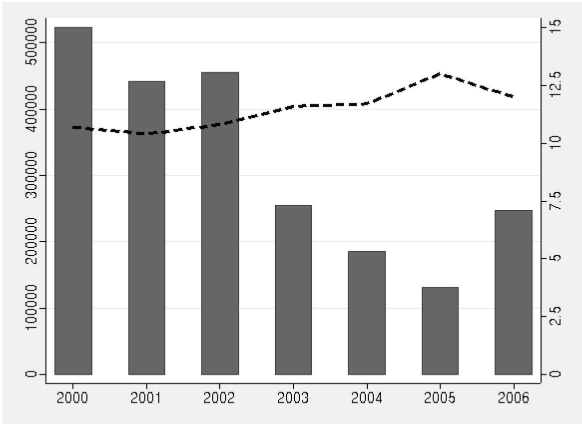
The remainder of this chapter is organized as follows. Section 3.2 describes the institutional background of public training programs in Germany with a particular focus on changes between the pre- and post-reform period. After describing the analytical framework in Section 3.3 and the data in Section 3.4, we discuss the matching quality in Section 3.5 and present the results in Section 3.6. Finally, Section 3.7 concludes.

## 3.2 Institutional Background

ALMP aims to increase the employment prospects of unemployed individuals. For this purpose, the Federal Employment Agency (FEA) in Germany spends a substantial amount of money on programs 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 3.1). While more than 500,000 unemployed individuals entered a training program in 2000, this number approached only around 130,000 individuals in 2005. In 2006, it increased again to nearly 250,000 persons entering such programs.

The effectiveness of public training programs in Germany before the Hartz reforms has been subject to a number of studies. For a recent review of the results

Figure 3.1: Entrants into public training programs, unemployment rate (2000–2006)




*Source:* Federal Employment Agency (FEA).
*Note:* Bars show annual number of entrants into public training programs (left axis).
The dashed line represents the average unemployment rate (right axis, in percent).

see Caliendo and Steiner (2005).[26] The results are quite heterogeneous—depending on the investigation period and the underlying data set. Recent studies which are based on rich administrative data sets often find at least for some sub-groups positive treatment effects (Lechner et al., 2005, 2007; Fitzenberger et al., 2008; Biewen et al., 2007; Rinne et al., 2007). However, there are also recent studies finding insignificant or negative effects (Hujer et al., 2006; Lechner and Wunsch, 2008). Besides differences in the investigation period and the underlying data set, 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-participants. Overall, the major lesson from the evaluation studies analyzing the pre-reform period—i.e., before 2003—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.

Germany's ALMP has undergone a series of reforms during the past years. Figure 3.2 summarizes the most important legislative changes in this context. Although these reforms are commonly referred to as the Hartz reforms, the first effort was made when the *JobAQTIVE Law* came into force on January 1, 2002. In addition to changes

[26]The international literature on the evaluation of ALMP is summarized by LaLonde (2003) and Kluve (2006), among others.

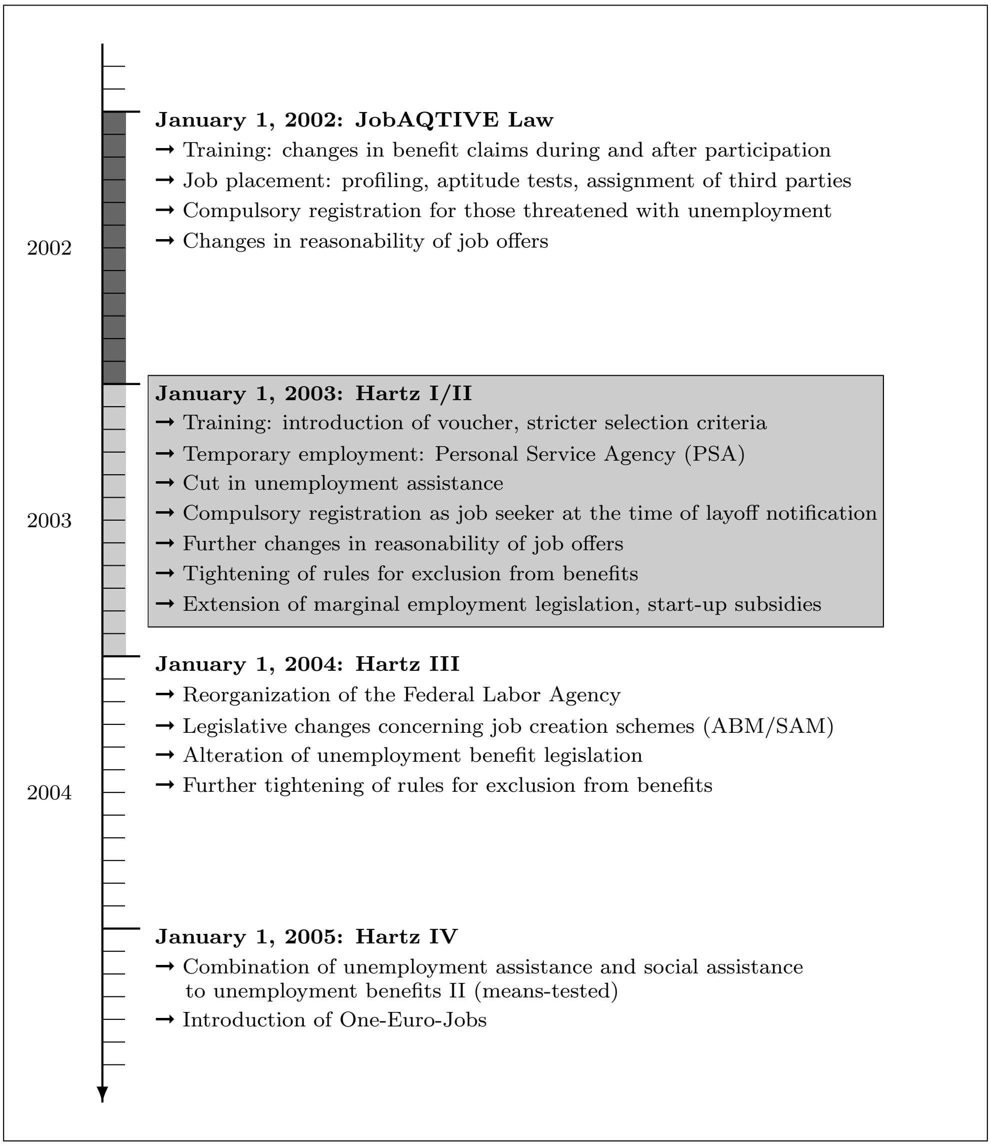
which affected job placement procedures, registration standards and job offer requirements, this law also altered the rules for benefit claims during and after participation in public training programs.

Major changes which affected the provision of public training programs came into force on January 1, 2003 under the first two reform packages (Hartz I/II). Additional changes affected temporary employment as Personal Service Agencies (PSA) were introduced and moreover unemployment assistance levels were cut. Registration standards and job offer requirements were altered (again), and the rules for exclusion from benefits were tightened. Finally, the legislation concerning marginal employment was extended and a new start-up subsidy for the unemployed was introduced.

The third package of Hartz reforms (Hartz III) came into force on January 1, 2004. Its main objective was a reorganization of the Federal Labor Agency. Additional changes affected job creation schemes, the unemployment benefit legislation, and the rules for exclusion from benefits. As a final step of the series of reforms, the fourth package (Hartz IV) was introduced in the beginning of 2005. Its most important feature constitutes the combination of the former systems of unemployment assistance and social assistance into the new means-tested unemployment benefit II system. Besides that, One-Euro-Jobs were introduced. An overview of the impact of the reform on ALMP is given for example by Jacobi and Kluve (2007). A broader picture is provided by Ebbinghaus and Eichhorst (2009) who describe institutional provisions and reforms of employment protection, active and passive labor market policies in Germany between 1991 and 2005.

Prior to the Hartz I/II package, i.e., before 2003, 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 limited since approvals were granted only to

Figure 3.2: Chronology of the Hartz reforms




*Source:* Authors' illustration.

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, 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.

After January 1, 2003 the provision of public training programs substantially changed. The most prominent feature of the reform marks the introduction of the training voucher (*Bildungsgutschein*) which abandoned the former *de facto* contracting-out system. 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 counterfactual 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. As Figure 3.3 shows, 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 at most three months. Within this period, job seekers are completely free to chose among approved training providers and courses in the market—subject to the requirements stated in the voucher.[27] 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).[28]

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

[27] The 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).

[28] The official transitional arrangement was as follows: "Individuals who were counseled 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 on or before February 28, 2003."

Figure 3.3: Training Voucher

**Training Voucher Number.: ________/____**
**Customer Number: ____________**
**(In accordance with § 77 para. 3 of the German Social Code III)**

Valid until:

Costs are covered
☐ according to the certification procedure (lump-sum)
☐ according to actual costs (as evidence is provided)

Maximum duration: up to ____ months (including an internship if necessary)

Educational target / qualifications provided:

Type of course: ☐ full-time (35 hours/week) ☐ part-time (12–24 hours/week)
☐ on-the-job course ☐ correspondence course

Provider: ☐ in-firm training ☐ off-the-job training

Location: ☐ outside a daily commuting distance

*Source:* Authors' illustration.

instance, information asymmetries constrain consumer sovereignty. In particular low-skilled job seekers lack the abilities to navigate the training market and to take an active role in searching for an appropriate course. This argument is supported by Kruppe (2008) who finds that low-skilled individuals are significantly less likely to redeem a granted voucher than persons with higher qualifications. However, the overall redemption rate is comparatively high with 85 percent in the period from 2003 to 2006 (Kruppe, 2008). On the supply side, a potential obstacle for competition among providers is their unequal distribution across German regions. Providers also reacted to the reform and increased co-operation and collusive behavior, for example by not offering 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.

Given the practical obstacles and the actual implementation process, the overall impact of the reform on the effectiveness of public training programs is anything but clear. However, Schneider and Uhlendorff (2006) and Schneider et al. (2007) 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.

Compared to previous studies, the most significant difference of our chapter is the decomposition of the overall reform effect into a "selection effect" and a "voucher effect". We refer to the selection effect as the effect resulting from a different composition of participants between the pre- and the post-reform period. This effect is due both to stricter selection criteria and to the *unintended* consequence of the voucher that low-skilled job seekers lack the abilities to navigate the training market and to take an active role in searching for an appropriate course. Low-skilled individuals thus exhibit a lower voucher receipt rate (intended by the caseworkers) as well as a lower voucher redemption rate (unintended). On the other hand, the voucher effect comprises the *intended* impacts of the introduction of training vouchers according to our taxonomy. These consequences include a potentially better match between participants and courses, an apparently more market-oriented (i.e., demand-oriented) approach of the local FEA offices, and quality enhancements which could be due to increased competition among training providers.

In order to isolate the above two effects, and to avoid complications of other components of Hartz III and IV reform discussed previously, we restrict our sample to 2002 (pre-reform period) and 2003 (post-reform period).

## 3.3 Analytical Framework

In order to disentangle the effects of the reform arising from a change of program quality due to a better match between participants and providers and/or an improved quality of the offered training programs (voucher effect) and the change of the composition of participants (selection effect), we apply a two-step matching approach.

Using the potential outcome framework as in Neyman (1923), Roy (1951), or 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 an indicator for participation, we can define different treatment effects in a similar way as Heckman and Vytlacil (1999, 2005):

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

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

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

In this chapter, we are interested in the treatment effect on the treated and its change induced by the reform. $R_i$ is an indicator which takes on the value 0 if we observe an individual before the reform and 1 if we observe an individual after the reform. The average treatment effects on the treated before and after the reform are given by:

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

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

A simple comparison of treated and non-treated individuals may be biased if participants and non-participants differ with respect to characteristics having an impact on the outcome $Y$. If treatment assignment is *strongly ignorable*, i.e. if selection is based on observable characteristics $X$ (conditional independence) and if observable characteristics of participants and non-participants overlap, the matching approach is an appealing choice to estimate treatment effects. This implies that unobserved characteristics may play a role for the selection into training, but that these unobserved characteristics have to be uncorrelated with the outcome variables, once we condition on $X$. Formally, these assumptions are given by:

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

$$0 < prob(D_i = 1|X_i) < 1 \qquad \text{(Overlap assumption)}$$

Rosenbaum and Rubin (1983) show that if the matching assumptions hold, i.e., treatment assignment is strongly ignorable *given* $X$, it is also strongly ignorable *given any balancing score that is a function of* $X$.[29] One possible balancing score is the propensity score $P(X)$, i.e., the probability of participating in a given program. Mueser et al. (2007) present evidence that if rich administrative data is used to measure the performance of training programs, propensity score matching is generally most effective.

We thus estimate $ATT_b$ ($ATT_p$) from pre-reform data (post-reform data) by propensity score matching methods.[30] However, the difference between $ATT_b$ and $ATT_p$ 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). If we assume additive separability of the two components, $ATT_p$ is given by:

$$ATT_p = ATT_b + VE + SE \tag{3.4}$$

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

$$\begin{aligned} RE &= ATT_p - ATT_b \\ &= VE + SE \end{aligned} \tag{3.5}$$

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. Note that we have a large sample of pre-reform participants. This implies that we will find for nearly all of our post-reform participants a corresponding

[29]When there are many covariates, it is impractical to match directly on covariates because of the curse of dimensionality. See, e.g., Zhao (2008) for some comments on this problem.

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

match, which ensures that our matched sample of participants is representative for the 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 $ATT_{bp}$ is the effect *only* for those participants under the pre-reform regime who are comparable to participants after the reform. This step controls for the changes in the composition of participants before and after the reform, i.e., the selection effect.

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

$$VE = ATT_p - ATT_{bp} \tag{3.6}$$

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

$$\begin{aligned} SE &= RE - VE \\ &= (ATT_p - ATT_b) - (ATT_p - ATT_{bp}) \\ &= ATT_{bp} - ATT_b \end{aligned} \tag{3.7}$$

There are several propensity score matching methods suggested in the literature, see, e.g., Imbens (2004), Caliendo and Kopeinig (2008) and Imbens and Wooldridge (2008) for overviews. Based on the characteristics of our data and particularly because of the two-step matching approach which is pursued, we opt for nearest-neighbor matching without replacement. This matching method has the advantage of being the most straightforward matching estimator: a given participant is matched with a non-participant or participant who is closest in terms of the estimated propensity score. We avoid an increased variance of the estimator as we match without replacement (Smith and Todd, 2005a), which is justified since the ratio between participants and (non-)participants—i.e., potential matching partners—is comparatively high in our data. Hence, the constructed counterfactual outcome is based only on distinct (non-)participants.

More specifically, the probability of participation is estimated conditional on a number of observable characteristics using binary probit models with participation as the dependent variable.[31] These characteristics include socio-demographic-characteristics (e.g., age, nationality, marital status), regional information (region, unemployment rate), educational and vocational attainment, the employment history (four years prior to program entry), and information on the last employment spell (duration, income, business sector).[32] We run these regressions separately for women and men from East and West Germany, respectively. After estimating the propensity score, we find a suitable matching partner by exact covariate matching combined with propensity score matching. The variables used for exact matching are region, previous unemployment duration (in months), and quarter of program entry. Therefore, we stratify the four sub-samples of women and men in East and West Germany by these variables first, and then implement propensity score matching for each cell without replacement. This procedure ensures that matched participants and non-participants are *a*) previously unemployed for the same duration at program entry, and *b*) entering the program in the same quarter. The latter condition makes sure that seasonal influences are constant. Furthermore, we do not condition on future non-partition. This is important in the context of dynamic assignment processes. Following the argumentation of Sianesi (2004), in countries like Sweden or Germany, in principle any unemployed individual will join a program at some time, provided he remains unemployed long enough. Hence, a restriction on future outcomes—i.e., to require non-participation in the follow-up period after the fictitious program entry—is likely to effect estimated treatment effects negatively, since a substantial fraction of the 'never treated'-individuals would *de facto* be observed to leave the unemployment register.

In order to assess the impact of the reform on the employment probability and earnings of participants, we estimate linear probability and ordinary least squares models. To test the robustness of our results with respect to potential differences in observable characteristics $X$, which may remain after the matching, we run additional

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

[32]Selected variables are listed in Table 3.1. The exact specifications and a list of all variables are not reported here, but available upon request.

regressions controlling for observable characteristics $X$.[33]

These additional regressions also allow us to control for changes in the general economic situation and changes in the extent and the composition of ALMP; these changes may be additional components of the reform effect.[34] Although we generally argue that we control for such changes, as participants and matched non-participants are subject to the same environment, we will explicitly address this issue in our sensitivity analyses and control for potential changes in our regressions.

For the variance of the estimated treatment effects, we base our inference on bootstrapping procedures. More specifically, we bootstrap the whole estimation process. This allows us to calculate standard errors based on the distribution of the estimated treatment effects. The standard errors of the reform effect, the voucher effect and the selection effect are based on the distribution of the respective differences in treatment effects across the replications.

## 3.4 Data

We use a sample of a particularly rich administrative data set, the Integrated Employment Biographies (IEB) of the FEA.[35] It contains detailed daily information on employment subject to social security contribution including occupational and sectoral information, receipt of transfer payments during periods of unemployment, 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.

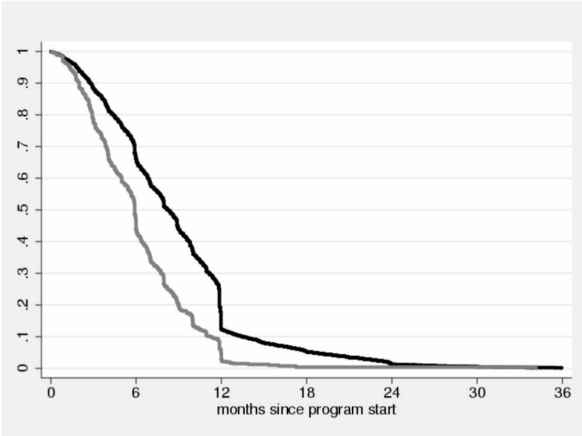
In Germany, public training programs for the unemployed are quite heterogenous. Thus, we concentrate on the most important program type: occupation-related or

[33] This procedure can correct a possible remaining bias as discussed in Rubin (1973), Imbens (2004), and Abadie and Imbens (2002).

[34] For instance, Lechner and Wunsch (2006b) present evidence for a clear positive relation between the effectiveness of the programs and the unemployment rate over time.

[35] The IEB combines 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). However, it 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., Jacobebbinghaus and Seth (2007) for details on the IEBS.

Figure 3.4: Actual program duration




*Source:* IEB, own calculations.

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

general training. Participants either 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). The program does *not* aim to provide a certificate, i.e., an officially recognized vocational degree. In contrast to other program types it focuses on classroom training and is neither provided in combination with internships nor is the simulation of real operations conducted. In the pre-reform period about 60 percent of all participants in public training programs were assigned to this particular type, and it became even more important after the reform in 2003 as this share increased to more than 70 percent. Figure 3.4 indicates that the program is—in comparison to other ALMP measures in Germany—a rather short measure. Both in the pre-reform and in the post-reform period, after one year more than 90 percent of the participants have exited from the program. However, the program duration decreased after the reform. While the median program duration is about 8 months in the pre-reform period, it amounts to about 6 months after the reform.

To evaluate the impact of the reform and its features on the effectiveness of this type of public training program, our data includes participants as well as non-participants from the pre- and post-reform period, respectively. More specifically, we have information on: *a*) participants who entered the program in 2002, *b*) participants who entered the program in 2003, *c*) non-participants in 2002, and *d*) non-participants in 2003. We do not have information on individuals who received a voucher but did not make use of it. Our sample of participants who entered the program in 2003 consists of more than 1,200 individuals. In order to apply the matching approach as described above (see Section 3.3) roughly 20 participants from the period before the reform were drawn per participant in 2003. Therefore, we have information on about 23,000 participants who entered the program in 2002. Beyond the matching of post-reform participants with pre-reform participants, we need to match participants with non-participants. For both years (2002 and 2003) our sample of non-participants—i.e., potential controls—consists of more than 500,000 individuals. Non-participants are required to not have participated in the given type of training before and in the quarter of the participant's program entry, but we do not condition on future non-participation.

Table 3.1 displays descriptive statistics of selected variables for the samples of participants and non-participants in 2002 and 2003, respectively. Focusing firstly on individuals participating in training, we find evidence for a change in their composition between the pre- and the post-reform period in our data. The most remarkable change can be observed with respect to the previous employment histories which differ considerably between these two groups. Considering a period of four years prior to program entry, participants who entered after the reform show a higher labor market attachment in terms of un-/employment rates and a slightly higher income from last employment than earlier program entrants. The average age of a participant also dropped by more than one year, while other characteristics remain on average rather stable between the two years. In particular, differences with respect to the educational or vocational attainment do not appear to be substantial. On the other hand, the groups of non-participants are very different from the groups of participants in both years. They are on average older and less educated. Moreover, their employment histories reveal a higher incidence of unemployment as well as a lower incidence of employment relative to participants in training.

The success of program participation is evaluated by looking at *a*) the probability of being employed, and *b*) earnings. Our observation period—i.e., the period in which outcomes are observed—starts at program entry and it ranges over a period of

Table 3.1: Descriptive statistics (selected variables)

| | Participants 2002 | | Participants 2003 | | Non-participants 2002 | | Non-participants 2003 | |
|---|---|---|---|---|---|---|---|---|
| | Mean | (Std. Dev.) | Mean | (Std. Dev.) | Mean | (Std. Dev.) | Mean | (Std. Dev.) |
| *Socio-demographic characteristics* | | | | | | | | |
| Female | 0.4156 | (0.4928) | 0.4224 | (0.4941) | 0.4271 | (0.4947) | 0.4293 | (0.4950) |
| East | 0.4298 | (0.4951) | 0.4086 | (0.4918) | 0.3029 | (0.4595) | 0.3115 | (0.4631) |
| Age | 37.248 | (9.5597) | 35.871 | (9.7840) | 39.975 | (11.439) | 39.547 | (11.143) |
| German citizenship | 0.9179 | (0.2745) | 0.9236 | (0.2657) | 0.8693 | (0.3370) | 0.8829 | (0.3215) |
| *Educational attainment* | | | | | | | | |
| No graduation | 0.0545 | (0.2270) | 0.0601 | (0.2378) | 0.1147 | (0.3186) | 0.1140 | (0.3179) |
| First stage of secondary level | 0.3263 | (0.4689) | 0.2965 | (0.4569) | 0.4922 | (0.4999) | 0.4722 | (0.4992) |
| Second stage of secondary level | 0.4207 | (0.4937) | 0.4452 | (0.4972) | 0.2879 | (0.4528) | 0.3033 | (0.4597) |
| Advanced technical college entrance qualification | 0.0664 | (0.2490) | 0.0772 | (0.2670) | 0.0365 | (0.1875) | 0.0399 | (0.1957) |
| General qualification for university entrance | 0.1321 | (0.3386) | 0.1210 | (0.3263) | 0.0687 | (0.2530) | 0.0705 | (0.2561) |
| *Vocational attainment* | | | | | | | | |
| No vocational degree | 0.1808 | (0.3848) | 0.1901 | (0.3925) | 0.3389 | (0.4733) | 0.3230 | (0.4676) |
| In-plant training | 0.6253 | (0.4841) | 0.6182 | (0.4860) | 0.5519 | (0.4973) | 0.5555 | (0.4969) |
| Off-the-job training, vocational school, technical school | 0.0953 | (0.2936) | 0.0983 | (0.2978) | 0.0659 | (0.2481) | 0.0727 | (0.2597) |
| University, advanced technical college | 0.0986 | (0.2982) | 0.0934 | (0.2911) | 0.0433 | (0.2034) | 0.0487 | (0.2153) |
| *Unemployment history* | | | | | | | | |
| Share of unemployment in 1st year before program entry | 0.5942 | (0.3080) | 0.5972 | (0.3059) | 0.6225 | (0.3192) | 0.6320 | (0.3175) |
| Share of unemployment in 2nd year before program entry | 0.2479 | (0.3293) | 0.2332 | (0.3209) | 0.3384 | (0.3688) | 0.3476 | (0.3740) |
| Share of unemployment in 3rd year before program entry | 0.2065 | (0.3164) | 0.1613 | (0.2770) | 0.3016 | (0.3743) | 0.2894 | (0.3669) |
| Share of unemployment in 4th year before program entry | 0.1851 | (0.3032) | 0.1462 | (0.2676) | 0.2715 | (0.3643) | 0.2636 | (0.3608) |
| *Employment history* | | | | | | | | |
| Share of employment in 1st year before program entry | 0.2687 | (0.3057) | 0.2560 | (0.2953) | 0.2260 | (0.2984) | 0.2087 | (0.2867) |
| Share of employment in 2nd year before program entry | 0.5176 | (0.4276) | 0.5604 | (0.4190) | 0.4181 | (0.4186) | 0.4208 | (0.4198) |
| Share of employment in 3rd year before program entry | 0.5287 | (0.4380) | 0.5897 | (0.4279) | 0.4438 | (0.4329) | 0.4648 | (0.4314) |
| Share of employment in 4th year before program entry | 0.5264 | (0.4379) | 0.5840 | (0.4282) | 0.4568 | (0.4344) | 0.4888 | (0.4303) |
| ln(Last income from employment) | 3.8728 | (0.6220) | 3.8825 | (0.6179) | 3.7950 | (0.6383) | 3.7901 | (0.6377) |
| # Observations | 22,839 | | 1,231 | | 520,483 | | 532,893 | |

*Source:* IEB, own calculations.

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 states until December 31, 2004. 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 this outcome measure. Moreover, the administrative data set only includes employment that is subject to social security contributions.[36] Additionally, we evaluate the effect of program participation on monthly earnings in the primary labor market. We apply the described definition of employment and consider remunerations associated with these spells in terms of monthly earnings.

In order to control for changes in the general economic situation which may constitute another component of the reform effect, we consider a number of economic and labor market characteristics available for each labor market district. We use monthly information on the share of unemployed, the share of vacancies, the share of participants in various ALMP measures (including public training programs) as well as GDP growth rates.[37] Table 3.3 in the Appendix reports the change of these variables between 2002 and 2004. For example, the unemployment rate slightly increased on average from around 10 percent in 2002 to around 10.7 percent in 2004, while the share of unemployed individuals participating in training programs decreased during this period.

Furthermore, the implementation of the reform may have varied across local FEA districts. We address this issue by using information about the subjective judgement of the Hartz reforms by administrators. This is obtained through a survey conducted in the beginning of 2005 in the management departments of the local FEA districts. The respondents are asked about the change of the job placement, the benefit granting, the administrative effort and the co-operation with third parties like training providers and employers, and the subjective judgment is on average rather positive. However, we observe heterogeneity in the judgements, and we will control for this in our regressions. The included items are reported in Table 3.4 in the Appendix.

[36] This means that, e.g., we do not observe self-employment.

[37] We include annual GDP growth rates for the 16 federal states since more disaggregated data is not available.

## 3.5 Matching Quality

We apply different strategies to evaluate the balancing of observable characteristics between the different groups after the matching.

One way to assess the matching quality is to compare the standardized bias before matching, $SB^b$, to the standardized bias after matching, $SB^a$. The standardized biases are defined as

$$SB^b = \frac{(\overline{X}_1 - \overline{X}_0)}{\sqrt{0.5 \cdot (V_1(X) + V_0(X))}} \;\; ; \;\; SB^a = \frac{(\overline{X}_{1M} - \overline{X}_{0M})}{\sqrt{0.5 \cdot (V_{1M}(X) + V_{0M}(X))}} \; , \qquad (3.8)$$

where $X_1$ ($V_1$) is the mean (variance) in the treated group before matching and $X_0$ ($V_0$) the analogue for the comparison group. $X_{1M}$ ($V_{1M}$) and $X_{0M}$ ($V_{0M}$) are the corresponding values after matching (Rosenbaum and Rubin, 1985). The mean standardized bias should be reduced after matching.

Following the suggestion of Sianesi (2004) we also re-estimate the propensity score on the matched sample to compute the pseudo-$R^2$ before and after matching. The pseudo-$R^2$ indicates how well the observable characteristics $X$ explain the probability of being treated. After the matching the pseudo-$R^2$ should be low because there should be no systematic differences between the treated and not treated individuals.

In a third approach we test the balancing following a suggestion by Smith and Todd (2005b) and estimate the following regression for each observable characteristic $x$ included in our preferred specification:

$$\begin{aligned} x_k = {} & \beta_0 + \beta_1 \widehat{PS(X)} + \beta_2 \widehat{PS(X)}^2 + \beta_3 \widehat{PS(X)}^3 \\ & + \alpha_0 D + \alpha_1 D\widehat{PS(X)} + \alpha_2 D\widehat{PS(X)}^2 + \alpha_3 D\widehat{PS(X)}^3 + \varepsilon_k \end{aligned} \qquad (3.9)$$

$D$ is the treatment indicator, $\widehat{PS(X)}$ the estimated propensity score, and $x_k$ is the observable characteristic $k$. For each $x$ we perform an $F$-test of the joint null hypothesis that all coefficients on terms involving $D$ equal zero. If the balancing score satisfies the balancing condition, $D$ should not provide any information about $x_k$.

Table 3.2 summarizes the results from the three balancing tests. We tested the balancing for different sub-samples: women and men in East and West Germany, respectively. Altogether, we perform five matching procedures: *a*) pre-reform participants are matched with post-reform participants, *b*) *unmatched* pre-reform participants are matched with pre-reform non-participants, *c*) *unmatched* post-reform participants are matched with post-reform non-participants, *d*) *matched* pre-reform participants are matched with pre-reform non-participants, and *e*) *matched* post-reform participants are matched with post-reform non-participants. *Unmatched* and *matched* participants may differ because we do not find for every participant after the reform a suitable match from the period before the reform, i.e., the *matched* participants are a subset of the *unmatched* participants. Overall, the balancing of the different matching procedures is quite satisfactory: the percentage biases are apparently reduced. More specifically, mean standardized biases in the matched samples are—with one exception—noticeably smaller than in the unmatched samples and are are mostly below five percent after matching. Likewise, the pseudo-$R^2$ after matching are fairly low and decrease substantially compared to before matching. Moreover, in most of the matching procedures our third test indicates that $D$ does not provide any information about any observable characteristics. However, some of our matching procedures perform better than others. We get the worst performance for our matching of participants before the reform with participants after the reform, especially for females in East Germany—although the third test indicates no problems for the participant-participant matching for any of our sub-samples. Therefore, we will check the sensitivity of our results to the inclusion of observable characteristics in our regressions based on the matched samples, and we have to be careful in the interpretation of our results for females in East Germany.

As an additional check of the matching quality, we plot the fraction of individuals being employed in the primary labor market before and after matching for a period of four years prior to participation. This approach follows, e.g., Heckman and Hotz (1989) and Mueser et al. (2007). Figures 3.5 and 3.6 show that the one-step as well as the two-step matching procedures generate comparison groups with employment probabilities prior to participation which are pretty close to those of the treatment groups. Although they are still not identical, the substantial differences before matching disappear across all matched samples. Moreover, Figure 3.6(a) points out very clearly that—before matching—participants after the reform have more favorable employment histories than participants before the reform.

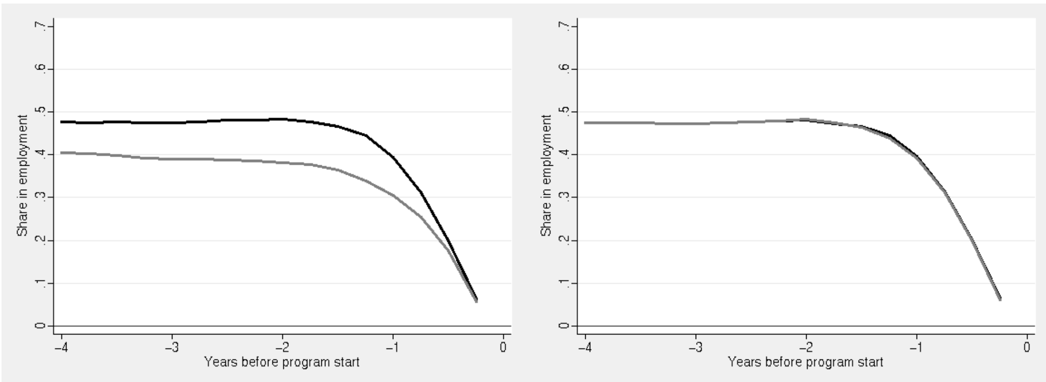
Table 3.2: Matching quality

| Sex | Region | *Before Matching* Mean bias | *Before Matching* $R^2$ | *After Matching* Mean bias | *After Matching* $R^2$ | # sign. |
|---|---|---|---|---|---|---|
| *a) Two-step matching: Participants 2002 – Participants 2003* | | | | | | |
| Female | East Germany | 8.359 | 0.2001 | 8.652 | 0.1329 | 0 |
| Female | West Germany | 10.211 | 0.2125 | 6.760 | 0.1054 | 0 |
| Male | East Germany | 7.493 | 0.0990 | 5.135 | 0.0595 | 0 |
| Male | West Germany | 8.184 | 0.1158 | 5.972 | 0.0769 | 0 |
| *b) One-step matching: Participants 2002 – Non-participants 2002* | | | | | | |
| Female | East Germany | 11.621 | 0.0536 | 1.059 | 0.0017 | 4 |
| Female | West Germany | 11.991 | 0.0693 | 0.982 | 0.0013 | 2 |
| Male | East Germany | 11.730 | 0.0562 | 1.050 | 0.0015 | 0 |
| Male | West Germany | 11.725 | 0.0584 | 0.895 | 0.0009 | 0 |
| *c) One-step matching: Participants 2003 – Non-participants 2003* | | | | | | |
| Female | East Germany | 13.461 | 0.0486 | 4.329 | 0.0337 | 0 |
| Female | West Germany | 15.383 | 0.0558 | 3.714 | 0.0278 | 0 |
| Male | East Germany | 14.231 | 0.0535 | 4.064 | 0.0271 | 1 |
| Male | West Germany | 13.358 | 0.0480 | 2.241 | 0.0106 | 0 |
| *d) Two-step matching: Matched Participants 2002 – Non-participants 2002* | | | | | | |
| Female | East Germany | 24.504 | 0.1092 | 5.811 | 0.0385 | 0 |
| Female | West Germany | 25.836 | 0.1068 | 3.402 | 0.0247 | 0 |
| Male | East Germany | 20.709 | 0.0881 | 2.969 | 0.0214 | 0 |
| Male | West Germany | 22.185 | 0.0724 | 2.629 | 0.0204 | 0 |
| *e) Two-step matching: Matched Participants 2003 – Non-participants 2003* | | | | | | |
| Female | East Germany | 21.841 | 0.0859 | 3.924 | 0.0333 | 0 |
| Female | West Germany | 23.526 | 0.0946 | 2.581 | 0.0121 | 0 |
| Male | East Germany | 18.842 | 0.0664 | 2.916 | 0.0219 | 0 |
| Male | West Germany | 18.242 | 0.0714 | 2.329 | 0.0151 | 0 |

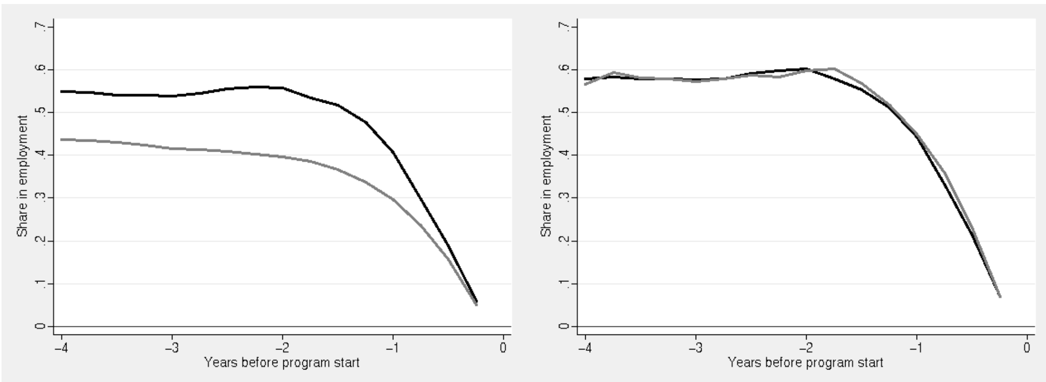
*Notes:* Mean Bias: Mean standardized bias; $R^2$: Pseudo-$R^2$ of propensity score estimation; # sign.: number observable characteristics for which $F$-test rejects the joint null. Further details are given in the text.

Figure 3.5: Pre-entry employment shares: one-step matching procedure

(a) unmatched P 2002 – unmatched NP 2002 (b) matched P 2002 – matched NP 2002

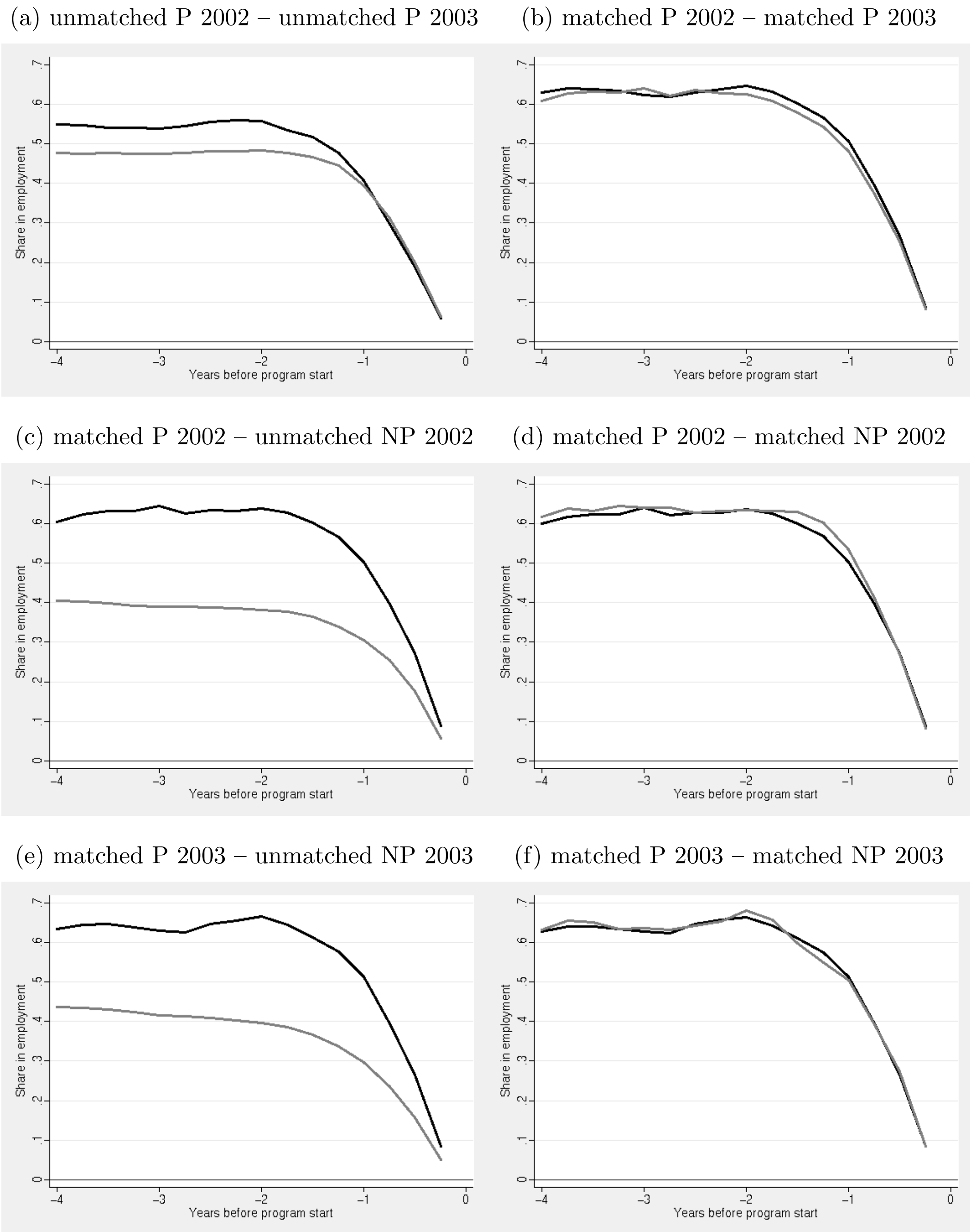



(c) unmatched P 2003 – unmatched NP 2003 (d) matched P 2003 – matched NP 2003




*Notes:* P: participants (in black); NP: non-participants (in gray).

Figure 3.6: Pre-entry employment shares: two-step matching procedure




(a) unmatched P 2002 – unmatched P 2003
(b) matched P 2002 – matched P 2003
(c) matched P 2002 – unmatched NP 2002
(d) matched P 2002 – matched NP 2002
(e) matched P 2003 – unmatched NP 2003
(f) matched P 2003 – matched NP 2003

*Notes:* P: participants (in black); NP: non-participants (in gray). Matching between participants 2002 and participants 2003: pre-reform participants in gray, post-reform participants in black.

## 3.6 Results

In this section, we present the effects on different outcomes for a period of 1.5 years after program entry. First, we compare the employment probabilities. Second, we present results for the effects on earnings. Additionally, we investigate whether the effects differ across skill-groups, and finally we conduct several sensitivity analyses.

### 3.6.1 Employment Probabilities

The estimates of the average treatment effects on employment probabilities of the participants are reported in Figure 3.7. We observe that participants before and after the reform face a substantial lock-in effect, both treatment effects are significantly negative in the first months.[38] After around 6 months of training, both treatment effects diverge and the treatment effect for participants after the reform constantly lies above the treatment for participants before the reform. At the end of our observation period, i.e., 1.5 years after program entry, the point estimates of the treatment effects amount to about 3 percentage points before the reform and about 7 percentage points after the reform. The differences between these two treatment effects describes the reform effect. We thus find a positive impact of the reform, which may occur due to the voucher effect or due to the change in the composition of participants.

Figure 3.8 displays the decomposition of the reform effect and reveals insights about the extent and magnitude of reform effect, voucher effect, and selection effect. The upper part reports the point estimates of the three effects, while the effects including the corresponding confidence intervals are reported separately in the lower part. The decomposition shows that the positive reform effect seems to be exclusively based on the voucher effect. Similar to the reform effect, the voucher effect becomes substantially positive after around 6 months. The voucher effect is significantly positive from month 7 until month 13 after entering the program. However, for the last 5 months of our observation period, the effect is still positive, but not significantly different from zero.

[38] While participating—or being 'locked-in' in the program—individuals probably reduce their search activities for new jobs (van Ours, 2004).

The point estimates of the selection effect are almost always negative, but never significantly different from zero. This indicates that there is no evidence for a positive impact of a stricter selection of participants on the average treatment effect. Our results suggest that the overall reform effect would have been more positive if the composition of participants had not changed. Our finding is consistent with Lechner and Smith (2007) who present evidence that caseworkers are not the best choice to allocate unemployed individuals into programs. Although their results are based on Swiss data, the situation in which caseworkers select the training providers and programs on behalf of the unemployed precisely describes the pre-reform situation in Germany. This changed under the new regime; and after the reform job seekers are free to choose their provider on their own by means of training vouchers.

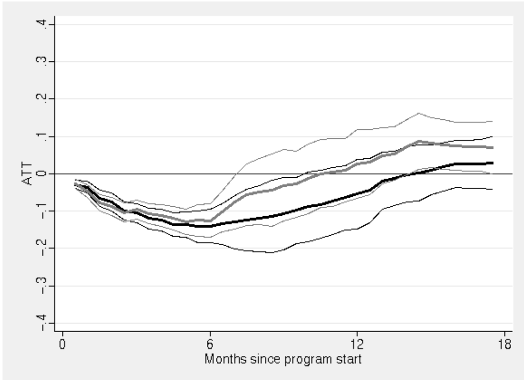
Additionally, we estimate the corresponding effects for four subgroups separately: men and women in West and East Germany, respectively. The results are reported in the Appendix, Figures 3.29–3.44. We find some evidence for heterogeneity of the effects, but the general picture is similar. We find rather negative and never significant selection effects and rather positive voucher effects. The voucher effect is significantly different from zero only for men in East Germany. However, because the number of observations clearly drops if we analyze the effects separately for the four subgroups, these results have to be interpreted with caution.

### 3.6.2 Earnings

Corresponding to the effects on employment probabilities, we present the average treatments effects on monthly earnings before and after the reform in Figure 3.9. Again, we observe substantial lock-in effects for both periods and clearly higher point estimates for the post-reform period after around 6 months of treatment. 18 months after entering the program, the point estimates of the treatment effects are about € 40 and roughly € 150 per month in the pre- and post-reform period, respectively.

Figure 3.10 displays the decomposition of the reform effect. Similar to the employment probabilities, the positive reform effect seems to be almost exclusively based on the voucher effect. We find no significant impact of the selection effect and a positive impact of the voucher effect, although not always significant. The similarity to

Figure 3.7: Reform effect, employment

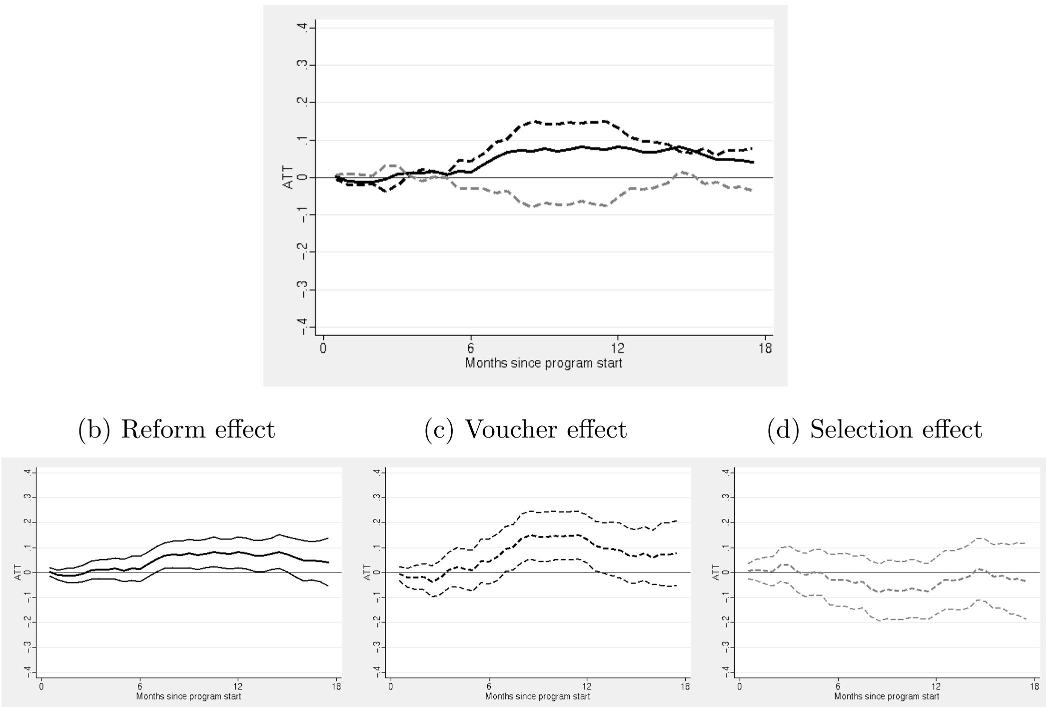



*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.8: Decomposition, employment

(a) Decomposition

(b) Reform effect (c) Voucher effect (d) Selection effect




*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

the employment probabilities is not surprising, because the positive earnings effects reflect at least partly the increased employment probabilities.

In order to get additional insights in the effects on earnings, we estimate the treatment effects using realized earnings only, i.e., we compare earnings conditional on being employed. The results are reported in Figures 3.11 and 3.12. The point estimates of the voucher effect are always positive, and the selection effect is always negative. Both effects are not significantly different from zero, but the voucher effect steadily increases during the observation period. Altogether, this indicates that the introduction of the voucher—next to an increased employment probability—also leads to better job matches for the participants, measured by on average higher monthly earnings in the new job.

### 3.6.3 Effects Across Skill Groups

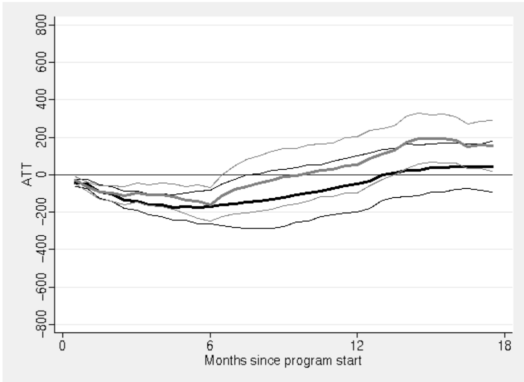
Preliminary evidence suggests that low-skilled job-seekers may lack the abilities to navigate the training market and to take an active role in searching for an appropriate course (Kruppe, 2008). If this is the case, the advantages of the introduction of vouchers will only partly occur in this group and mainly occur among skilled individuals. To assess this issue in more detail, we differentiate between two skill groups—skilled and unskilled individuals—and analyze the reform effects as well as the voucher and selection effects separately for those groups.

Our classification of skilled and unskilled individuals is based on whether or not an individual has received a formal vocational degree before entering the program.[39] This distinction closely follows Dustmann and Meghir (2005) who define skill groups similarly. The importance of this distinction is emphasized as the authors find substantial differences between the two groups in terms of job mobility, wage growth, and returns to experience. In their view, these differences have important implications, e.g., for the design of ALMP.[40]

[39]We consider completed in-plant training and off-the-job training as well as degrees from a vocational school, a technical school, a university, or a university of applied sciences as vocational degrees.

[40]For a broader and more general overview about about the German system of secondary school tracks, the apprenticeship system, and vocational degrees which can be obtained, see for example Winkelmann (1996) and Dustmann (2004).

Figure 3.9: Reform effect, earnings (definition A)

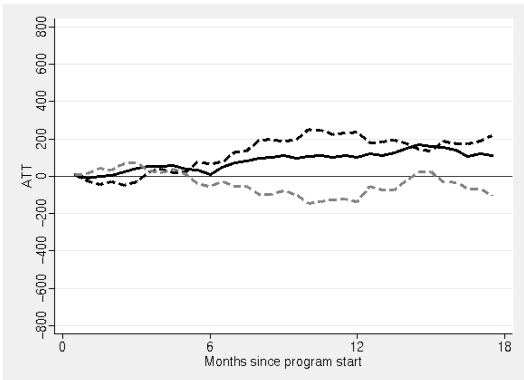



*Definition A:* Monthly earnings where no earnings are treated as zero.

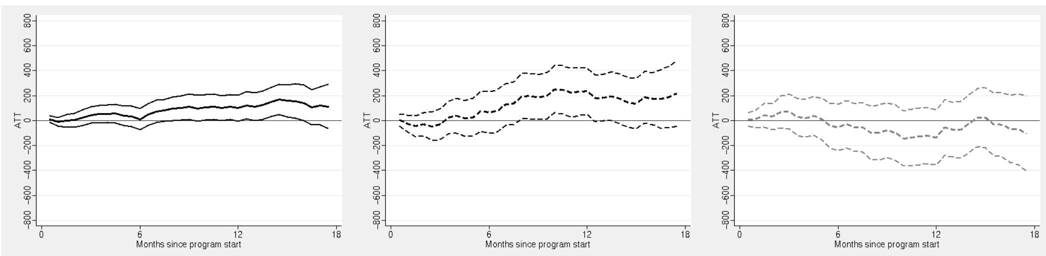
*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.10: Decomposition, earnings (definition A)

(a) Decomposition



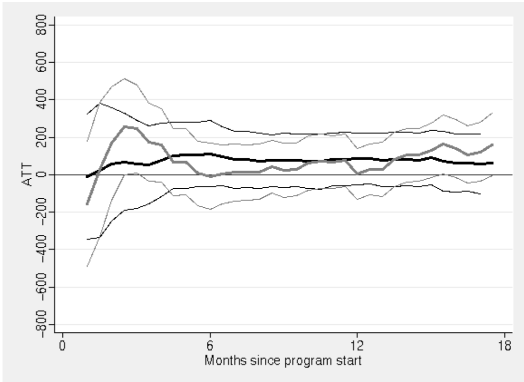

(b) Reform effect | (c) Voucher effect | (d) Selection effect




*Definition A:* Monthly earnings where no earnings are treated as zero.

*Note:* Total reform effect in black (solid), voucher effect in black (dashed), and selection effect in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.11: Reform effect, earnings (definition B)



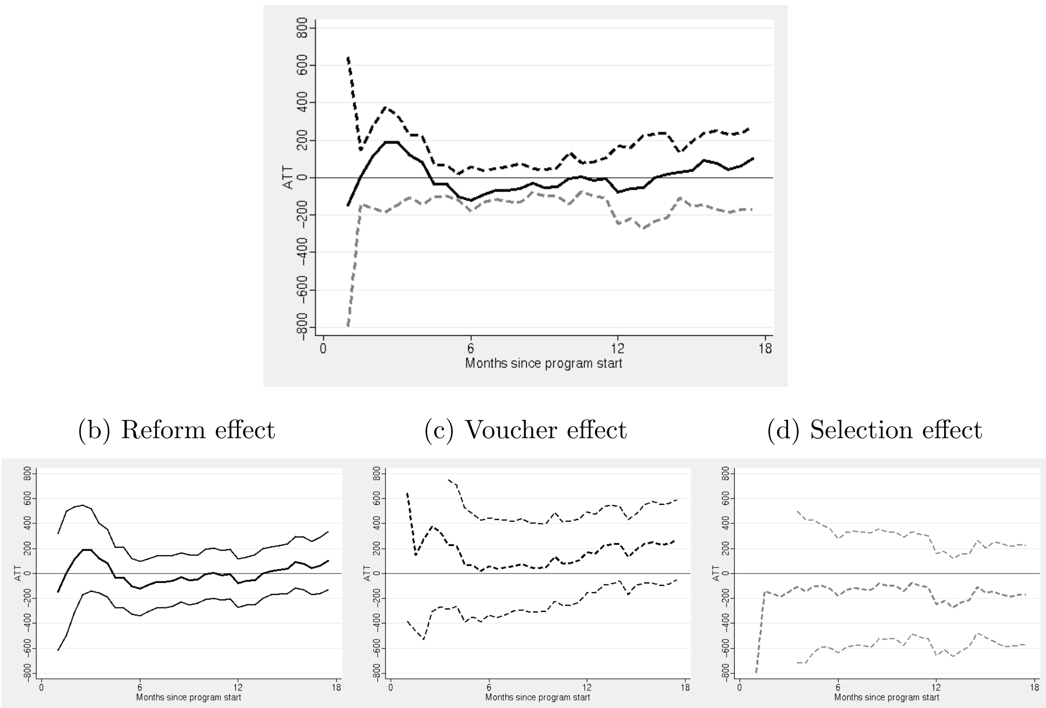

*Definition B:* Monthly earnings where no earnings are treated as missing.

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.12: Decomposition, earnings (definition B)

(a) Decomposition

(b) Reform effect (c) Voucher effect (d) Selection effect




*Definition B:* Monthly earnings where no earnings are treated as missing.

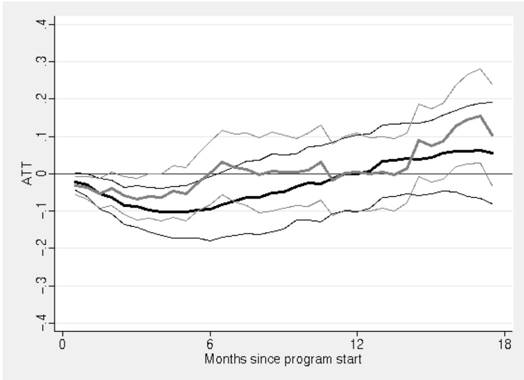
*Note:* Reform effect in black (solid), voucher effect in black (dashed), and selection effect in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figures 3.13 and 3.14 display the reform effect and the results of the decomposition with respect to employment probabilities for the subgroup of unskilled individuals. Indeed, we do not find any significant impacts of the reform. More specifically, although we find slightly positive overall reform effects on employment probabilities, these effects turn out not to be significantly different from zero during the observation period. Both voucher effects and selection effects are virtually zero for the group of unskilled. On the other hand, Figures 3.15 and 3.16 support the notion of both significantly positive overall reform effects and significantly positive voucher effects for skilled individuals. The selection effects are not significantly different from zero, but they are in general negative.

Figures 3.17–3.20 display the average treatments effects on monthly earnings before and after the reform as well as the decomposition of the reform effect for unskilled and skilled individuals. The impacts on monthly earnings are virtually the same as on employment probabilities. We again observe positive reform and voucher effects for skilled individuals, whereas the selection effect is—if at all—negative. On the other hand, we neither find significant overall reform effects nor significant voucher and selection effects for unskilled individuals.

Overall, our results for the two different skill groups suggest that both the positive reform effects and the positive voucher effects which we find in the whole sample only arise for skilled individuals. This group can take advantage from an increased consumer sovereignty, whereas unskilled individuals indeed seem to have problems in adequately using the newly introduced voucher. However, it should be noted that the treatment effects do not change between pre- and post-reform period. Therefore, unskilled individuals who participate in the program are not worse off after the reform, but the innovative voucher scheme also does not improve the effectiveness of public training program for this group. Whether unskilled job seekers are free to select the training provider in the market or caseworkers make this choice simply appears not to matter in terms of program effectiveness.

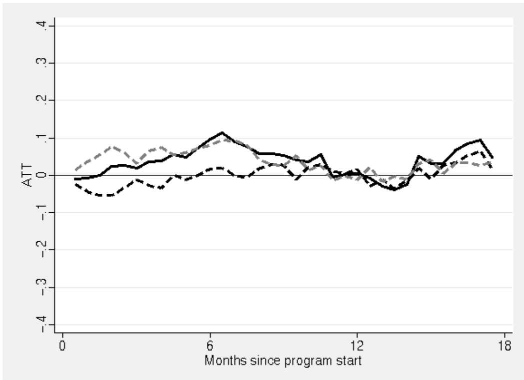
Figure 3.13: Reform effect, employment (unskilled)



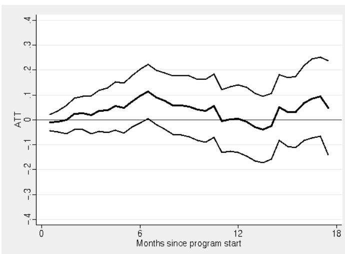

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.14: Decomposition, employment (unskilled)

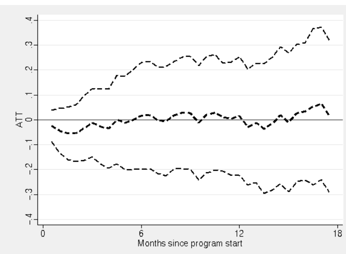
(a) Decomposition



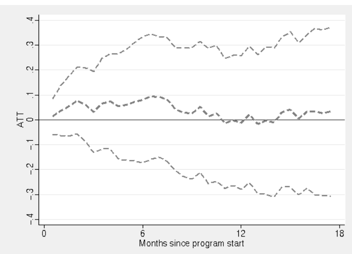

(b) Reform effect



(c) Voucher effect

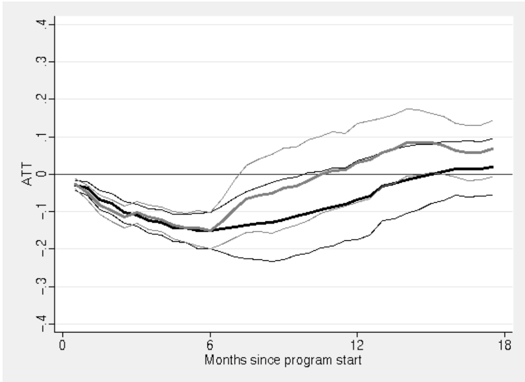


(d) Selection effect



*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.15: Reform effect, employment (skilled)

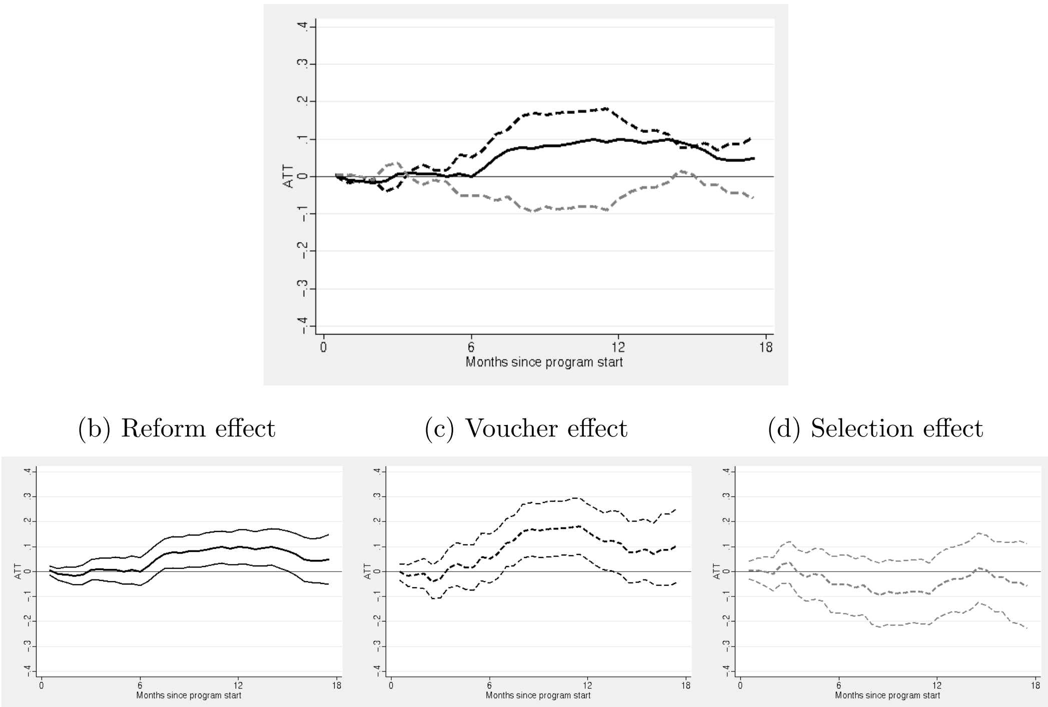



*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.16: Decomposition, employment (skilled)

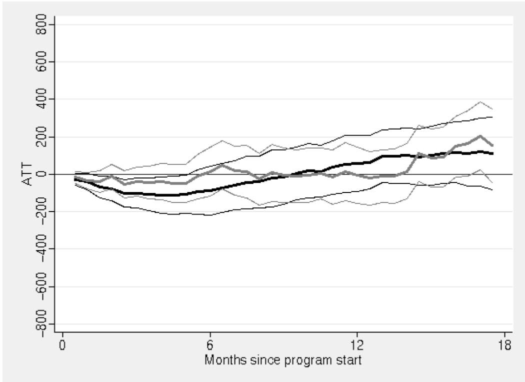
(a) Decomposition

(b) Reform effect (c) Voucher effect (d) Selection effect




*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.17: Reform effect, earnings (definition A, unskilled)

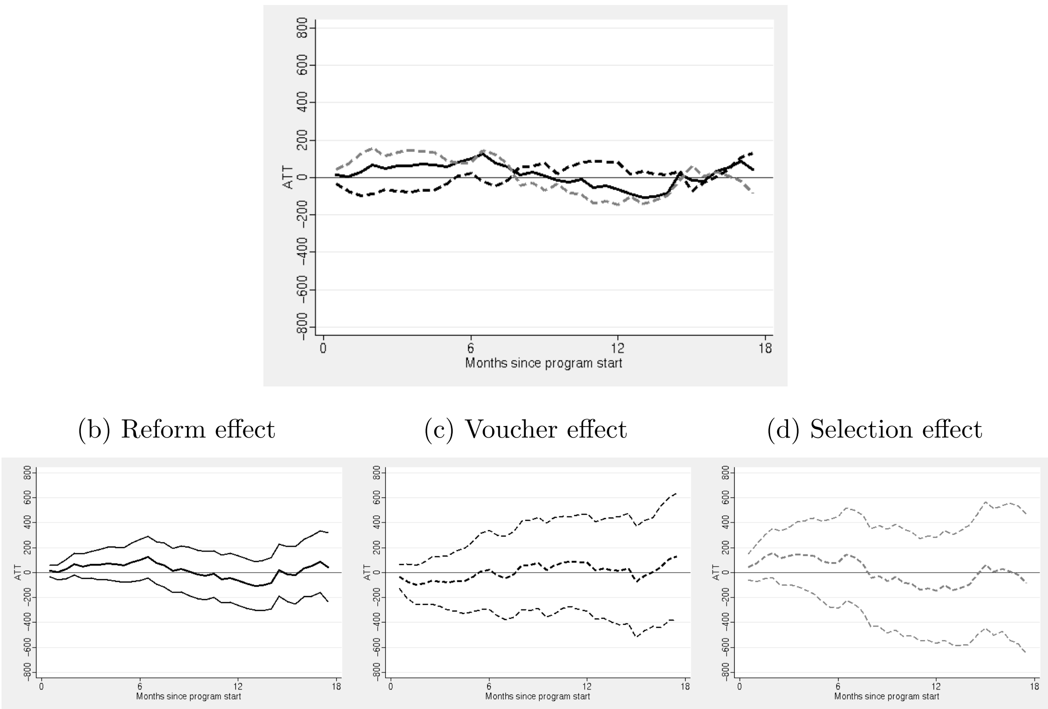



*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.18: Decomposition, earnings (definition A, unskilled)

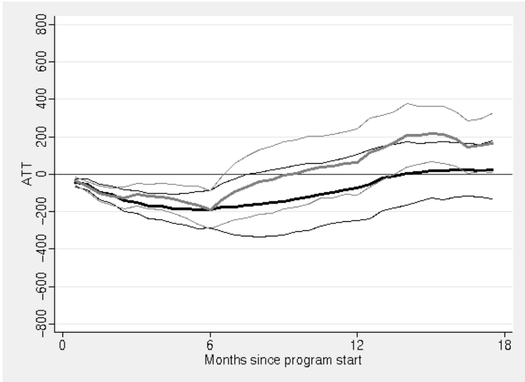
(a) Decomposition

(b) Reform effect (c) Voucher effect (d) Selection effect




*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

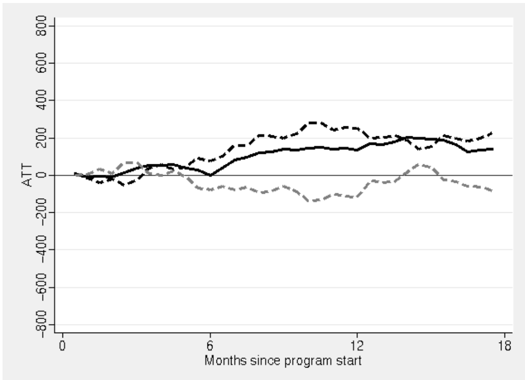
Figure 3.19: Reform effect, earnings (definition A, skilled)



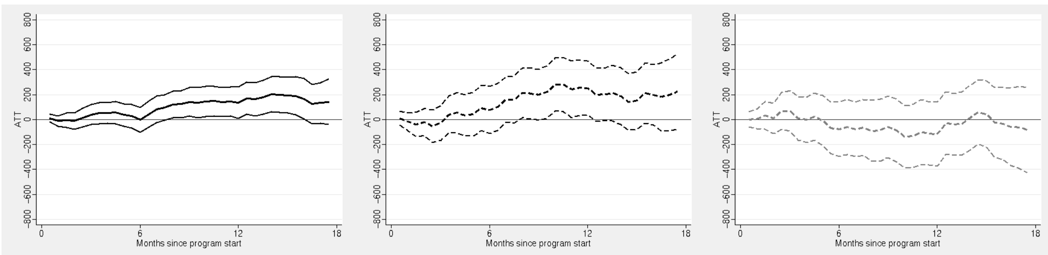

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.20: Decomposition, earnings (definition A, skilled)

(a) Decomposition




(b) Reform effect (c) Voucher effect (d) Selection effect




*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

### 3.6.4 Sensitivity Analysis

We address the robustness of our previous results in this section. For this purpose, we perform a sensitivity analysis in two steps. We assess the robustness of our results with respect to *a*) the inclusion of additional control variables, and *b*) the transitional arrangement for the training voucher in the beginning of 2003.

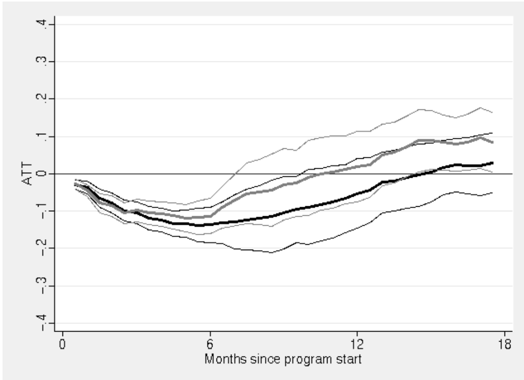
#### Inclusion of additional control variables

One may argue that changes in the general economic situation constitute another component of the reform effect. Therefore, we additionally control for a number of economic and labor market characteristics which are available for each local FEA district. These variables are changing over time. In addition to that, we include observable individual characteristics measured before entering the treatment and include—only for the post-reform period—indicators describing the implementation of the Hartz reform on the FEA district level.

The results are presented in Figures 3.21 and 3.22. In general, the picture is very similar to the results presented above. The point estimates of the voucher effect are slightly lowered, while the selection effect is slightly less negative. However, the voucher effect is still significantly positive between month 7 and month 13 after entering the program, and the selection effect is still almost always negative.

The results are also very similar for earnings, as reported in Figures 3.23 and 3.24. The voucher effect marginally lowers and selection affect slightly increases. Our results thus appear to be robust to the inclusion of additional control variables.

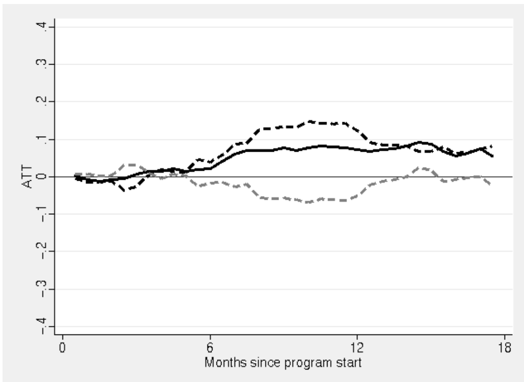
Figure 3.21: Reform effect, employment (including additional control variables)



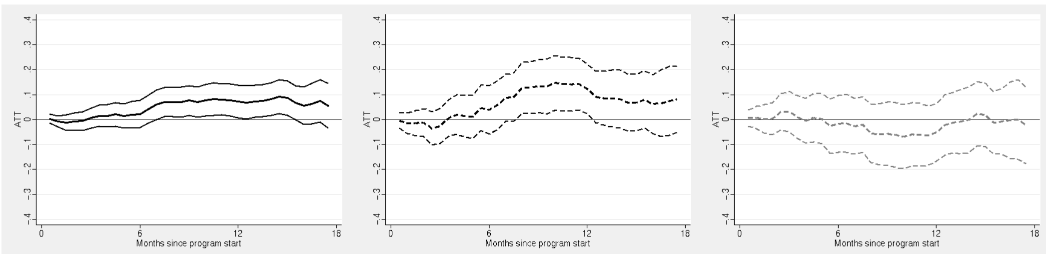

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.22: Decomposition, employment (including additional control variables)

(a) Decomposition

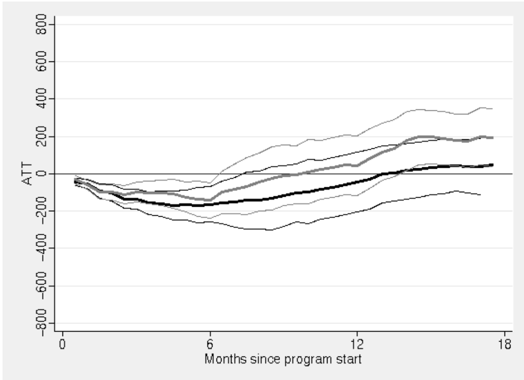



(b) Reform effect (c) Voucher effect (d) Selection effect




*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.23: Reform effect, earnings (definition A, including additional control variables)

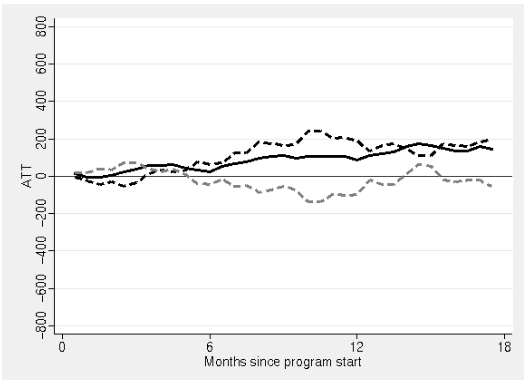



*Definition A:* Monthly earnings where no earnings are treated as zero.

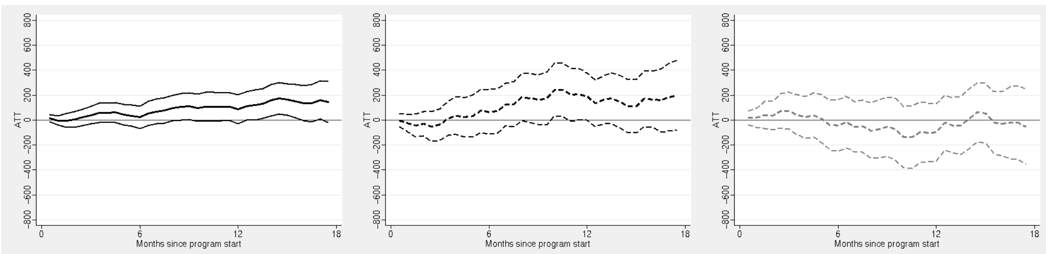
*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.24: Decomposition, earnings (definition A, including additional control variables)

(a) Decomposition




(b) Reform effect (c) Voucher effect (d) Selection effect



*Definition A:* Monthly earnings where no earnings are treated as zero.

*Note:* Total reform effect in black (solid), voucher effect in black (dashed), and selection effect in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

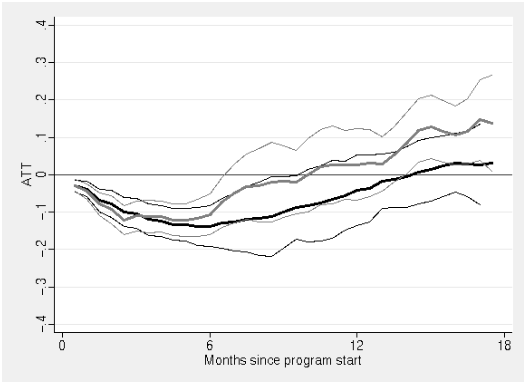
#### Transitional Arrangement

We mentioned above that there has been a transitional arrangement in place until March 2003 (see Section 3.2). Unfortunately, the administrative data set does not allow us to identify those participants who actually received and redeemed a training voucher. We thus perform a sensitivity analysis and exclude participants who entered public training programs in the first quarter of 2003.[41] The results of this analysis are depicted in Figures 3.25 and 3.26.

Also after taking account of the transitional arrangement in the beginning of 2003, we still observe the main result of a positive impact of the voucher. The selection effect is virtually zero. However, at the end of the observation period we estimate both voucher and selection effect to be positive, although not significantly different from zero. The results are again very similar for earnings, as reported in Figures 3.27 and 3.28.

[41] According to Schneider et al. (2007) who analyze survey data, the fraction of participants in public training programs actually receiving a voucher was about 30 percent in the first quarter of 2003, but sharply increased subsequently. Of course excluding participants who entered public training programs in the first quarter of 2003 implies that we also exclude participants who entered public training programs in the first quarter of 2002 as well as corresponding non-participants based on our matching algorithm.

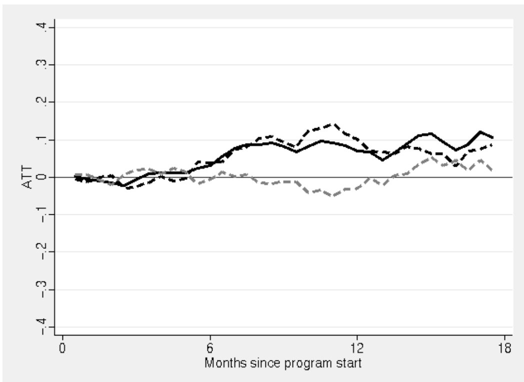
Figure 3.25: Reform effect, employment (excluding first quarter, including additional control variables)



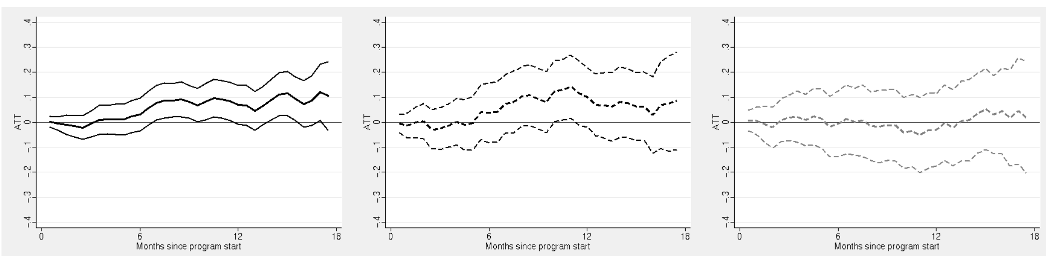

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.26: Decomposition, employment (excluding first quarter, including additional control variables)

(a) Decomposition

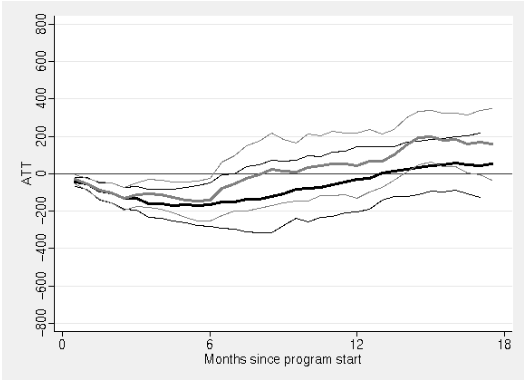



(b) Reform effect (c) Voucher effect (d) Selection effect




*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.27: Reform effect, earnings (excluding first quarter, definition A, including additional control variables)

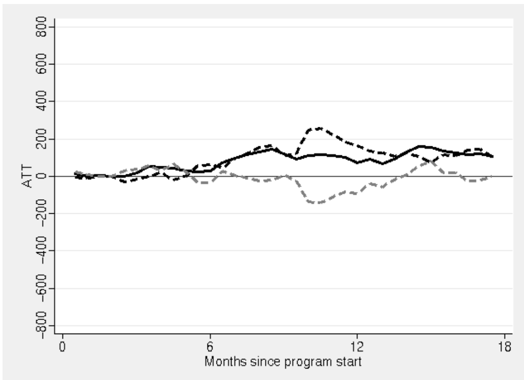



*Definition A:* Monthly earnings where no earnings are treated as zero.

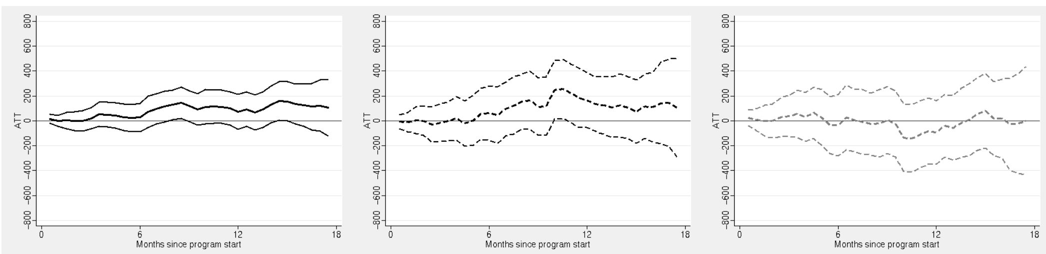
*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.28: Decomposition, earnings (excluding first quarter, definition A, including additional control variables)

(a) Decomposition




(b) Reform effect (c) Voucher effect (d) Selection effect




*Definition A:* Monthly earnings where no earnings are treated as zero.

*Note:* Total reform effect in black (solid), voucher effect in black (dashed), and selection effect in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

## 3.7 Conclusion

This chapter analyzes the impact of the labor market reform in 2003 on the effectiveness of the most important type of public training program in Germany. This reform had two main features: *a*) it introduced training vouchers, and *b*) it imposed more selective criteria on participants. Next to estimating the overall impact, we decompose the reform effect into *a*) the voucher effect, and *b*) the selection effect.

We find a slightly positive impact of the reform. The decomposition of this overall effect shows that the selection effect is—if at all—slightly negative. The voucher effect increases both the employment probability and earnings of the participants. This effect becomes substantially positive after around 6 months of training, and decreases slightly at the end of our observation period (1.5 years after program entry). Our results are mainly driven by skilled participants. We do not find any significant reform effect for the unskilled. While the former group can take advantage from an increased consumer sovereignty, unskilled individuals seem to have problems in adequately using the newly introduced voucher.

## 3.8 Appendix

Table 3.3: Economic and labor market variables

| | 2002<br>Mean<br>(Std. Dev.) | 2003<br>Mean<br>(Std. Dev.) | 2004<br>Mean<br>(Std. Dev.) |
|---|---|---|---|
| 1) Job seekers | 0.1238<br>(0.0630) | 0.1329<br>(0.0603) | 0.1441<br>(0.0617) |
| 2) Unemployment rate | 0.0991<br>(0.0499) | 0.1057<br>(0.0491) | 0.1065<br>(0.0490) |
| 3) Vacancies | 0.0107<br>(0.0050) | 0.0082<br>(0.0044) | 0.0065<br>(0.0038) |
| *Participants in ...* | | | |
| 4) Public training programs | 0.0085<br>(0.0052) | 0.0065<br>(0.0037) | 0.0045<br>(0.0023) |
| 5) Subsidized employment | 0.0037<br>(0.0043) | 0.0040<br>(0.0049) | 0.0029<br>(0.0037) |
| 6) Job creation schemes | 0.0050<br>(0.0076) | 0.0037<br>(0.0059) | 0.0029<br>(0.0047) |
| 7) GDP growth rate | 1.5191<br>(1.3170) | 1.0258<br>(0.6455) | 2.2837<br>(0.7575) |

*Source:* Federal Employment Agency (FEA); Statistical Offices of the Federal States.

*Notes:* 1)–6) are monthly shares in the civilian labor force in 178 FEA districts. 7) are annual GDP growth rates for the 16 Federal States.

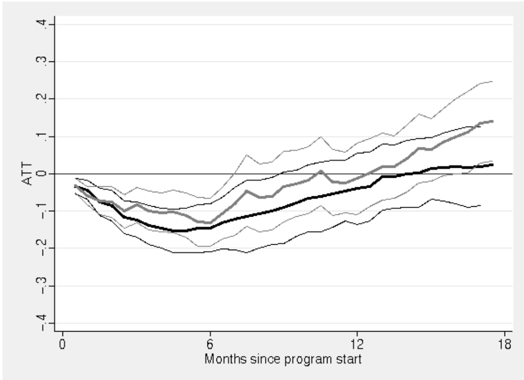
Table 3.4: Rating of the Hartz reforms by FEA districts

| *How did the reforms* | negative | | neutral | positive | | missing |
|---|---|---|---|---|---|---|
| *affect the ...* | –2 | –1 | 0 | +1 | +2 | . |
| ... effectiveness of job placement | 1 | 8 | 81 | 66 | 3 | 4 |
| | 0.6 | 4.9 | 49.7 | 40.5 | 1.8 | 2.5 |
| ... process of job placement | 1 | 9 | 61 | 82 | 6 | 4 |
| | 0.6 | 5.5 | 37.4 | 50.3 | 3.7 | 2.5 |
| ... efficiency of job placement | 1 | 7 | 64 | 83 | 5 | 3 |
| | 0.6 | 4.3 | 39.3 | 50.9 | 3.1 | 1.8 |
| ... process of benefit granting | 2 | 14 | 67 | 64 | 13 | 3 |
| | 1.2 | 8.6 | 41.1 | 39.3 | 8.0 | 1.8 |
| ... co-operation with third parties | 2 | 34 | 108 | 15 | 1 | 3 |
| | 1.2 | 20.9 | 66.3 | 9.2 | 0.6 | 1.8 |
| ... administration effort | 23 | 60 | 56 | 18 | 1 | 5 |
| | 14.1 | 36.8 | 34.4 | 11.0 | 0.6 | 3.1 |
| ... matching accuracy of job placement | 1 | 6 | 71 | 75 | 6 | 4 |
| | 0.6 | 3.7 | 43.6 | 46.0 | 3.7 | 2.5 |

*Source:* Survey in 163 FEA districts conducted in the beginning of 2005.

*Notes:* First row: frequencies; second row: percentages.

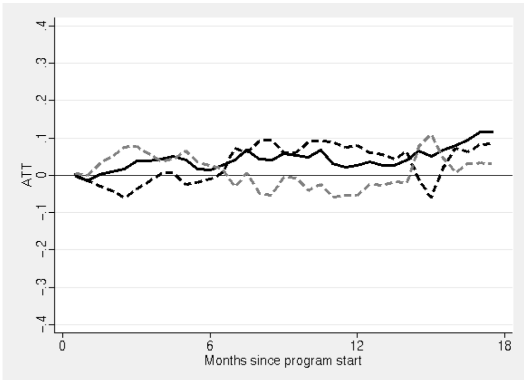
Figure 3.29: Reform effect, West German Men, employment



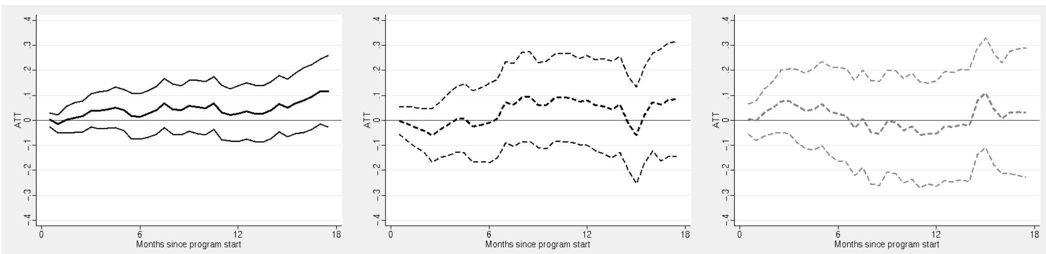

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.30: Decomposition, West German Men, employment

(a) Decomposition

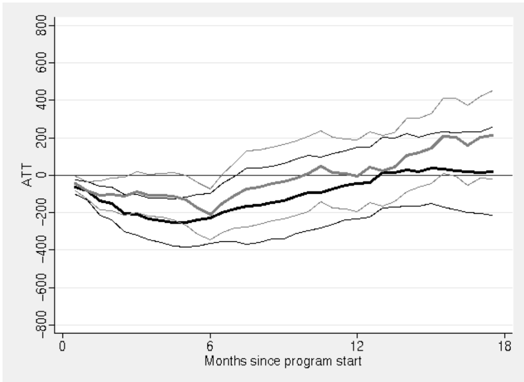



(b) Reform effect (c) Voucher effect (d) Selection effect




*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.31: Reform effect, West German Men, earnings (definition A)

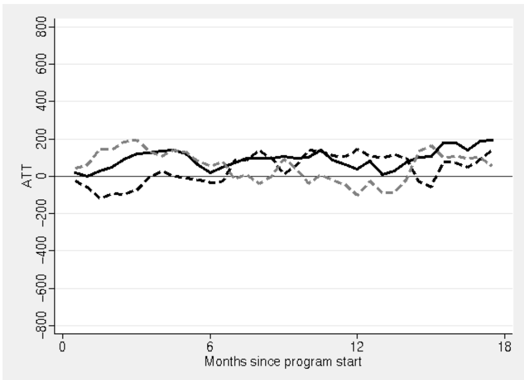



*Definition A:* Monthly earnings where no earnings are treated as zero.

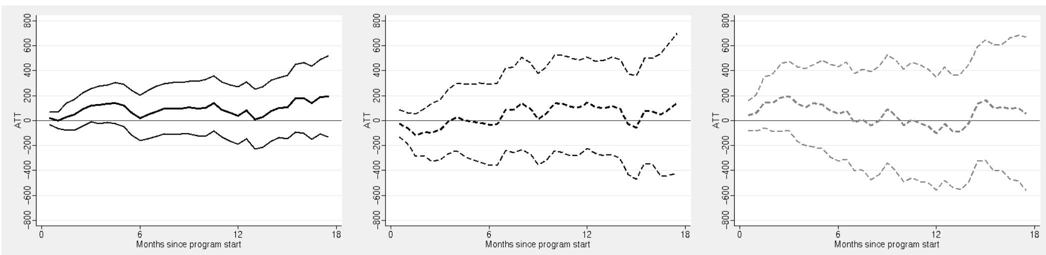
*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.32: Decomposition, West German Men, earnings (definition A)

(a) Decomposition



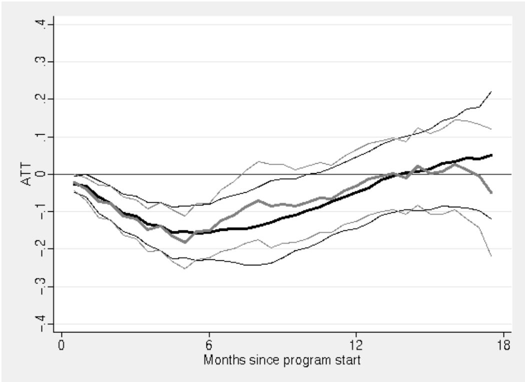

(b) Reform effect (c) Voucher effect (d) Selection effect




*Definition A:* Monthly earnings where no earnings are treated as zero.

*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

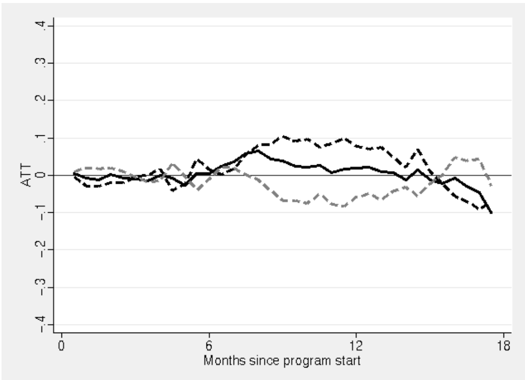
Figure 3.33: Reform effect, West German Women, employment



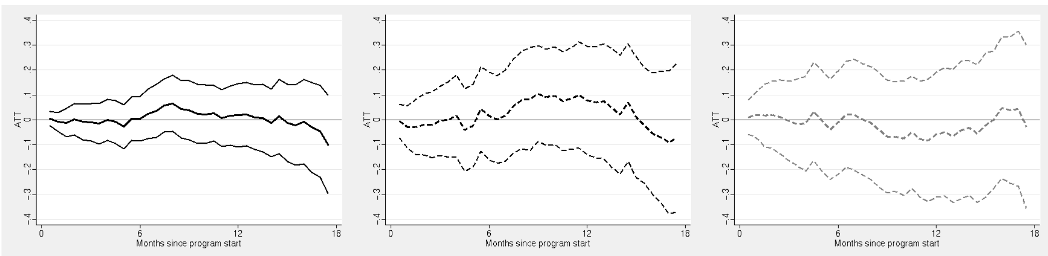

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.34: Decomposition, West German Women, employment

(a) Decomposition

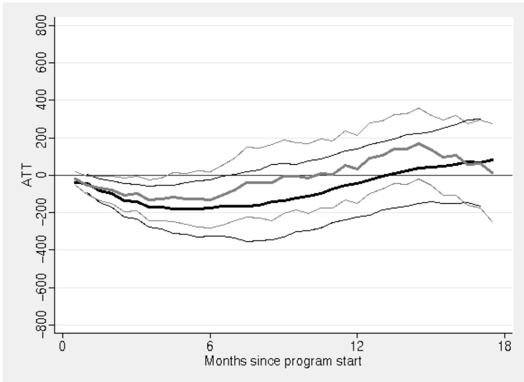



(b) Reform effect (c) Voucher effect (d) Selection effect




*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.35: Reform effect, West German Women, earnings (definition A)

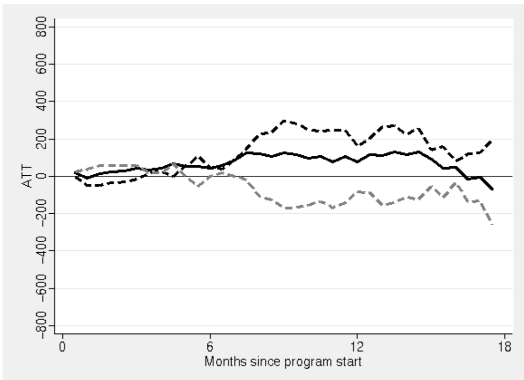



*Definition A:* Monthly earnings where no earnings are treated as zero.

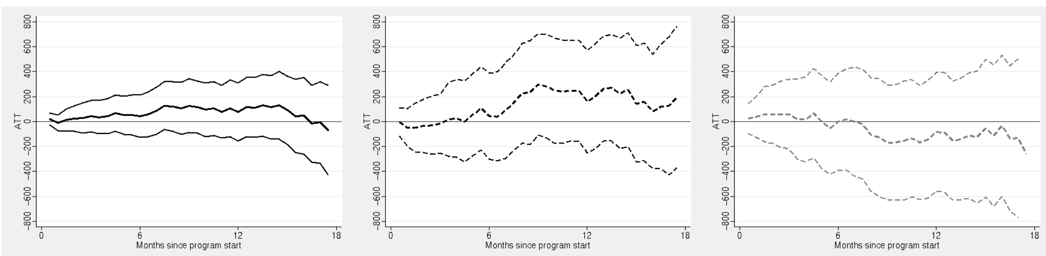
*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.36: Decomposition, West German Women, earnings (definition A)

(a) Decomposition



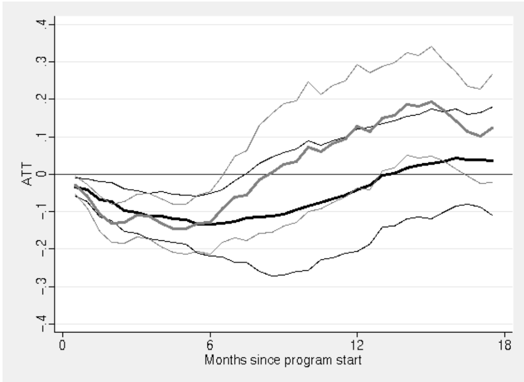

(b) Reform effect (c) Voucher effect (d) Selection effect



*Definition A:* Monthly earnings where no earnings are treated as zero.

*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

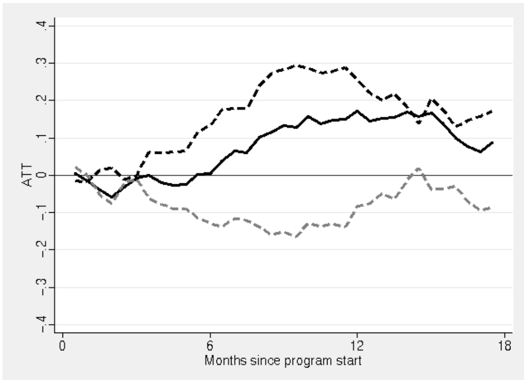
Figure 3.37: Reform effect, East German Men, employment



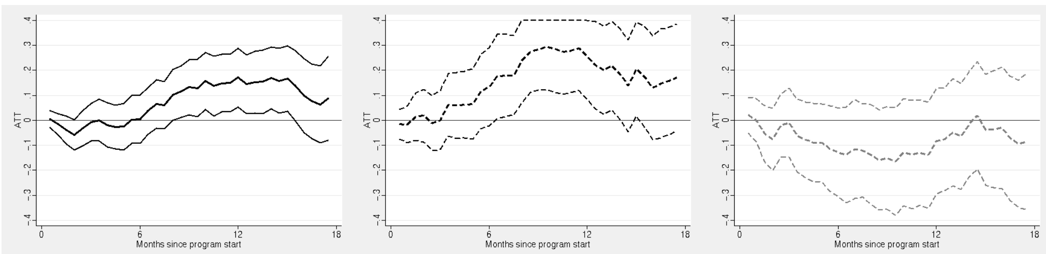

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.38: Decomposition, East German Men, employment

(a) Decomposition

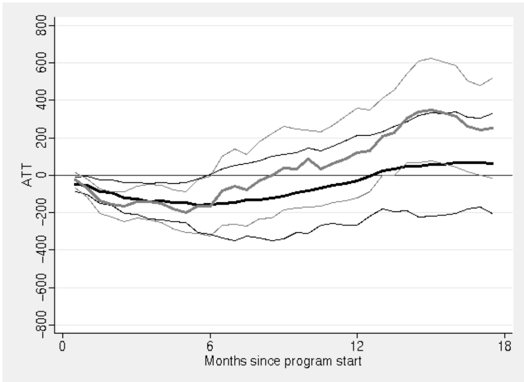



(b) Reform effect (c) Voucher effect (d) Selection effect



*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.39: Reform effect, East German Men, earnings (definition A)

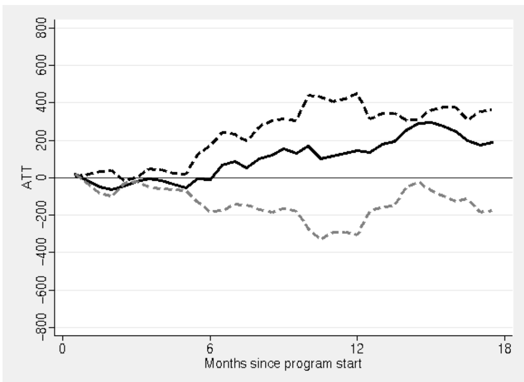



*Definition A:* Monthly earnings where no earnings are treated as zero.

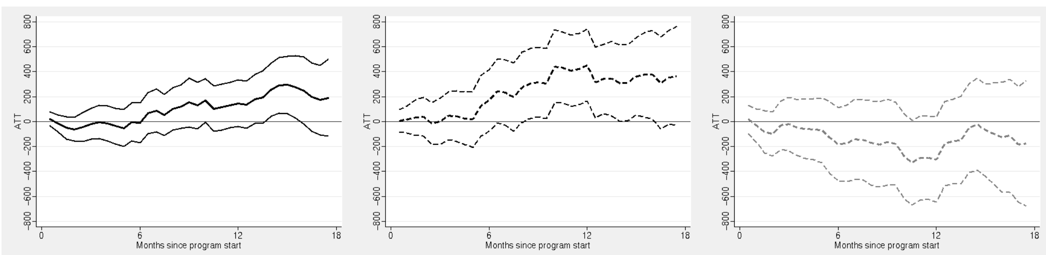
*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.40: Decomposition, East German Men, earnings (definition A)

(a) Decomposition



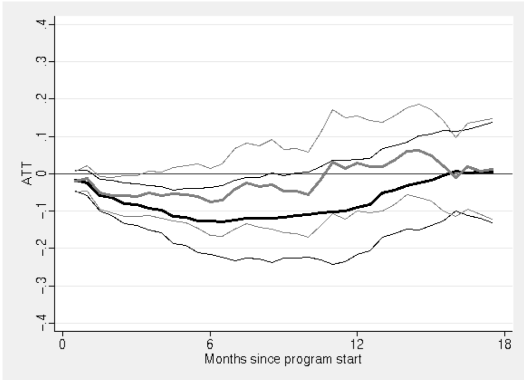

(b) Reform effect (c) Voucher effect (d) Selection effect




*Definition A:* Monthly earnings where no earnings are treated as zero.

*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

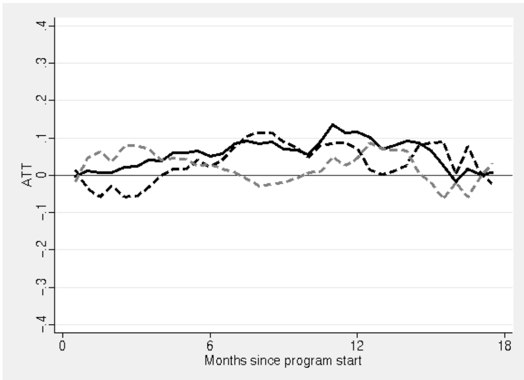
Figure 3.41: Reform effect, East German Women, employment



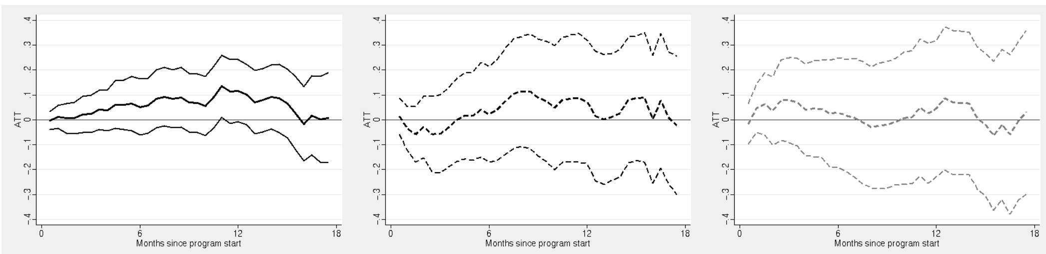

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.42: Decomposition, East German Women, employment

(a) Decomposition

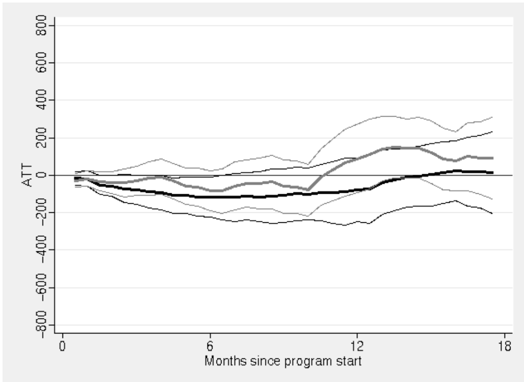



(b) Reform effect (c) Voucher effect (d) Selection effect




*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.43: Reform effect, East German Women, earnings (definition A)

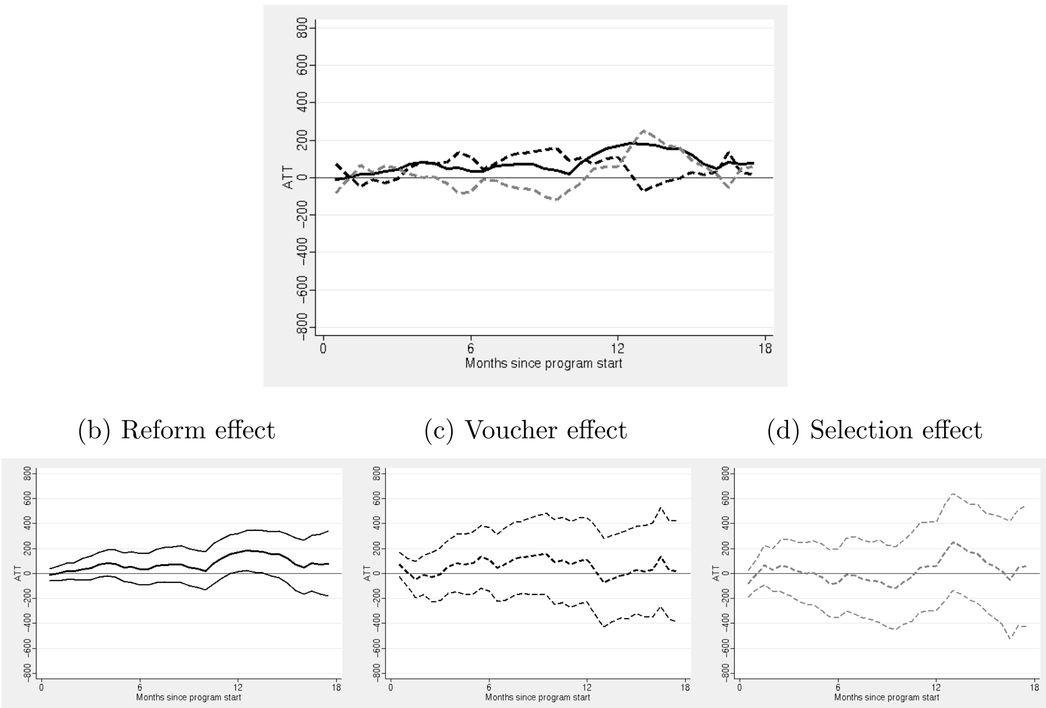



*Definition A:* Monthly earnings where no earnings are treated as zero.

*Note:* Pre-reform period in black, post-reform period in gray. Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

Figure 3.44: Decomposition, East German Women, earnings (definition A)

(a) Decomposition




(b) Reform effect

(c) Voucher effect

(d) Selection effect

*Definition A:* Monthly earnings where no earnings are treated as zero.

*Note:* Total reform effect (RE) in black (solid), voucher effect (VE) in black (dashed), and selection effect (SE) in gray (dashed). Thick lines refer to point estimates, thin lines indicate 95 percent confidence intervals.

# Chapter 4

# Active Labor Market Policy in a Transition Economy: Beautiful Serbia

*This chapter contributes to the still relatively scarce literature analyzing the effectiveness of active labor market policy in transition economies.*[42] *More specifically, the* Beautiful Serbia *program is analyzed. It consists of two basically independent parts, as it provides (a) training and (b) temporary work in the construction sector.*

[42]This chapter is based on joint work with Holger Bonin (Bonin and Rinne, 2006).

## 4.1 Introduction

This chapter studies the causal impact of participation in an active labor market policy (ALMP) measure—the *Beautiful Serbia* program providing training and temporary work in the construction sector in Serbia—on labor market outcomes as well as on subjective well-being approximating individual welfare. According to our estimates, the positive impact of this particular program appears much stronger judged by subjective well-being than judged by the immediate labor market effect.

Our study goes beyond the scope of traditional evaluation analysis which focuses on economic outcomes, i.e., judges the success of a labor market program by comparing the employment rates, unemployment rates, or wages of individuals who participate to the outcomes of comparable individuals who do not.[43] But a program may be beneficial for participants even if it does not immediately improve their labor market situation. It may reduce the psychic costs of being unemployed by strengthening participants' self-confidence or social contacts, and thus improve the subjective level of well-being. Lechner and Wunsch (2006a) give specific examples for different spheres—other than earnings and employment in the primary labor market—treatment effects can materialize in. They mention the facts of receiving earnings from work instead of benefits and having a daily routine in this context. More generally, in the economic literature on happiness a variety of measures of subjective well-being frequently serve as proxies for individual welfare.[44]

Nonetheless, so far the literature evaluating specific policies with respect to their impact on individual well-being is rather scarce. For instance, Gruber and Mullainathan (2005) assess the impact of a higher tax on cigarettes on the happiness of smokers, Di Tella et al. (2003) look at the impact of changes in unemployment benefits, and Frey and Stutzer (2000) analyze the role of direct democracy for subjective well-being. In the context of ALMP, Korpi (1997) shows that program participants indicate a higher level of subjective well-being than the openly unemployed who do

[43]See, e.g., Heckman et al. (1999), Martin and Grubb (2001) and Kluve (2006) for surveys of the international literature on the evaluation of ALMP measures.

[44]Frey and Stutzer (2002), Clark et al. (2006, 2008), and Di Tella and MacCulloch (2006) review the literature on happiness. Hayo and Seifert (2003) and Hayo (2004) show that most of the findings known from studies on the U.S. or Western Europe carry over to transition economies.

not attend such a program. However, our study can still be considered as one of the first that explicitly incorporates subjective well-being into the evaluation of ALMP.

This chapter also contributes to the still relatively small literature analyzing the effectiveness of ALMP in transition economies. Papers evaluating labor market programs in Eastern Europe—with rather mixed results—include studies focusing on Poland (Puhani and Steiner, 1997; O'Leary, 1998b; Kluve et al., 1999; Puhani, 2002), Romania (Rodriguez-Planas and Benus, 2009; Rodriguez-Planas, 2007), Slovakia (Lubyova and van Ours, 1999), Hungary (O'Leary, 1998a; Micklewright and Nagy, 2005), and Estonia (Leetmaa and Võrk, 2004). In this context, our paper provides the first evaluation of an ALMP measure implemented in Serbia.

The remainder of this chapter is organized as follows: Section 4.2 describes the background and design of the *Beautiful Serbia* program. Section 4.3 discusses our data. After explaining the evaluation strategy in Section 4.4, program impacts on labor market outcomes and subjective well-being are quantified in Section 4.5. Finally, Section 4.6 concludes.

## 4.2 The *Beautiful Serbia* Program

The Serbian economy is still considered to pass through a transitional phase. Although the country has initiated a package of economic reforms aimed at restructuring and liberalizing the economy and some positive results already materialized, high unemployment is very persistent. This is supposed to be partly an inherited problem and partly due to the prolonged and until 2000 highly irregular transition process. But also after the democratic changes of October 2000—the fall of the Milošević regime—unemployment has further increased. Table 4.1 displays the economic development of Serbia between 2000 and 2006. Although the economy has been improving in terms of real GDP and GDP per capita during this period, this process has surprisingly not translated into greater employment. The share of unemployed individuals among the economically active population continuously increased from about 12 (27) percent in 2001 to roughly 22 (33) percent in 2006 according to the labor force survey (ad-

Table 4.1: Economic indicators in Serbia (2001–2006)

| | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 |
|---|---|---|---|---|---|---|
| GDP per capita[a] | 1757 | 2242 | 2408 | 2643 | 2833 | 3424 |
| GDP real growth | 5.1 | 4.5 | 2.4 | 9.3 | 6.3 | 5.7 |
| Unemployment rate (LFS)[b] | 12.2 | 13.3 | 14.6 | 18.5 | 21.8 | 21.6 |
| Unemployment rate (admin.)[c] | 26.8 | 29.0 | 31.7 | 31.6 | 32.4 | 33.2 |

*Source:* Arandarenko and Jovicic (2007), Table 1.

*Notes:* [a] in Euros at exchange rate; [b] LFS: Labor Force Survey, in percent, October; [c] in percent, end of period, excl. agricultural self employment.

ministrative data).[45] This picture becomes even more dramatic for an earlier period: Arandarenko (2004) reports that unemployment in Serbia was 73 percent higher in 2000 than in 1993. Therefore, the issue of active labor market programs as temporary measures to alleviate the unemployment impact of the economic transition process is ranked high on the political agenda in Serbia, at least until the conditions of a rapid and sustained economic growth—accompanied by increasing employment—are established. The program under study represents one of the first policies implemented in the country for this purpose.

The *Beautiful Serbia* program operated in 2004 and 2005. It was administered by the United Nations Development Program, UNDP. The program was implemented with the support of the Ministry of Labor, Employment and Social Policy, MLESP, and fully incorporated into the National Employment Service of Serbia.[46] Due to limited financial means, the program was run on a small scale. It first started operating only in the capital city of Belgrade. In a second stage, which took place mostly in 2005, the geographic focus of the program shifted to the major cities of Niš and Zrenjanin. These three cities are economically quite heterogeneous (see Table 4.2). Overall, Belgrade appears to be the economically more advantaged region both in terms of the unemployment rate and the GDP level. However, in all three cities the unemployment rate reached a peak a few years ago (2002 in Belgrade, 2003 in Niš and Zrenjanin) and decreased afterwards. Additionally, the GDP levels somehow

[45] The differences in unemployment rates between administrative data and data from the labor force survey are mainly due to different underlying definitions of employment, unemployment, and participation (Arandarenko and Jovicic, 2007).

[46] Besides UNDP and MLESP, also the governments of Canada, the Netherlands, Austria and Greece, as well as city beneficiaries contributed to financing the program.

Table 4.2: Regional economic indicators (2001–2005)

| | 2001 | 2002 | 2003 | 2004 | 2005 |
|---|---|---|---|---|---|
| *Unemployment rate by district* | | | | | |
| City of Belgrade | 20.3 | 24.7 | 22.9 | 21.1 | 18.9 |
| Nišavski (incl. Niš) | 36.5 | 38.9 | 41.9 | 38.2 | 32.4 |
| Srednje-banatski (incl. Zrenjanin) | 33.6 | 38.2 | 41.6 | 40.2 | 38.5 |
| *GDP level by city (Republic of Serbia = 100)* | | | | | |
| Belgrade | 157.23 | 164.26 | 152.00 | 142.50 | 119.40 |
| Niš | 121.46 | 119.02 | 105.89 | 114.80 | 109.50 |
| Zrenjanin | 111.50 | 116.11 | 91.98 | 110.40 | 126.90 |

*Source:* Arandarenko and Jovicic (2007), Table 2, and Statistical Office of the Republic of Serbia.

converged in recent years and are fairly equal in 2006.

The basic design of the *Beautiful Serbia* program was intended to replicate the *Beautiful Bulgaria* program, an active labor market and refurbishing program which had run on a nation-wide scale in Bulgaria. The apparent success of this earlier program led officials to the assumption that it could be adapted to successfully work also in Serbia. The *Beautiful Serbia* program consists of two different components: (a) provision of vocational training for disadvantaged unemployed persons, and (b) subsequent provision of temporary jobs restricted to the (any) disadvantaged unemployed. The two components of the program are basically independent. Participation in the vocational training stage is neither a necessary nor a sufficient condition for obtaining a job offer in the temporary employment stage.

To be specific, the vocational training measure of the Beautiful Serbia program lasts for three months and is full-time. It provides certified vocational training for the constructional sector as mason, carpenter or painter. Its intended target group consists of long-term and otherwise disadvantaged unemployed persons, identified as such by the National Employment Service. No sanctions are applied if a person refuses to participate, and participation in the training measure can be considered as voluntary. Participants in vocational training receive a compensation amounting to about 30 percent of the average national wage. As only a very small fraction of the job seekers in Serbia are entitled to income support, this appears like a substantial incentive to take up the program.[47] Nonetheless it turned out difficult to attract individuals to the vocational training stage. One possible explanation is that participants supposedly face substantial opportunity costs in terms of forgone wage earnings. A large share of the unemployed in Serbia actually make their living from informal activities.[48] As the vocational training in the *Beautiful Serbia* program is full-time, participation is difficult to reconcile with these activities. We would thus expect that only those individuals expecting to recover the opportunity costs of their investment into human capital (self-)select into the training measure.

The second component of the *Beautiful Serbia* program is the provision of temporary jobs in the construction sector. Typical for a transition economy at its first growth stage, construction still plays an important role in the Serbian economy. Naturally temporary project-based jobs show a high incidence in this sector. The *Beautiful Serbia* program creates additional demand for these jobs by financing the refurbishment of selected public buildings and spaces. In the refurbishment projects, private firms are contracted under the condition that they employ a specified share (40–60 percent) of workers who are identified by the National Employment Service as previously unemployed and otherwise disadvantaged. Firms receive a fixed payment for conducting the refurbishment project. Projects are assigned to firms on a competitive basis, i.e., the firm offering the best quality-price ratio wins the tender. This procedure should guarantee that wages paid on the jobs in the temporary employment stage of the Beautiful Serbia program are competitive. In particular, firms do not receive a special wage subsidy for hiring the mandatory number of previously unemployed workers. The contracted firms can select among the pool of people who meet the criteria of the National Employment Service. Individuals run through an ordinary application procedure. Hence one would expect that successful candidates are hired in accordance with the needs of the company, and represent the most competent and capable among the unemployed individuals firms can choose among. This means that it is neither necessarily the case that participants in the vocational training part of the *Beautiful Serbia* program later on work in the sponsored refurbishment projects, nor that the

[47]The participants who were entitled to any kind of income support before the training received 110 percent of this amount during the period of training.

[48]Schneider (2007) presents estimates for the size of the Serbian shadow economy. It amounts to 37.3 percent of the official GDP in 2004/05.

previously unemployed workers hired for these projects did participate in the training measure before.

In total the *Beautiful Serbia* program provided vocational training to 252 unemployed individuals. The drop-out rate at this stage was very low. Almost 95 percent of the enrolled completed the training. In the 35 refurbishment projects financed by the program—managed by 16 contracted private companies—321 men found temporary employment.[49] About half of them had participated in the vocational training stage before.

## 4.3 Data

Our data come from a special survey of 363 individuals who were registered as unemployed at the National Employment Service when the program started (January 2004) and who either participated in at least one stage of the *Beautiful Serbia* program or did not participate at all. The interviews were conducted face-to-face by a professional survey agency, GfK Belgrade, shortly after the final refurbishment project of the program had been completed, during October and November 2005.

In principle, the survey was constructed such as to mimic an experimental design *ex post*. For each individual who participated in the *Beautiful Serbia* program, a matched partner with the same observable characteristics was drawn from the unemployment registry and scheduled for interview. The intention was to create a control group, which would resemble the treatment group as much as possible, with a limited the number of interviews. Unfortunately, due to deficiencies of unemployment registries at the National Employment Service, only few individual characteristics were available to implement this strategy. In effect, the one-to-one pre-matching routine to create a control group was only based on the following individual characteristics: age, education, and place of residence (Belgrade, Niš, or Zrenjanin). In particular, the (un-)employment history which appears extremely relevant for the success of ac-

[49] In principle, the program was available for women, too, but actually none participated—neither in the vocational training nor in the temporary employment stage.

Table 4.3: Distribution of observations across participation status

| | Participation in temp. employment? | | |
|---|---|---|---|
| Participation in training? | No | Yes | Total |
| No | 146 | 28 | 174 |
| Yes | 48 | 66 | 114 |
| Total | 194 | 94 | 288 |

tive labor market polices could not be controlled for.[50] In the accomplished survey, systematic drop-outs may further reduce the effective quality of the matches between program participants and non-participants. A sizeable number of persons scheduled for interview—around 40 percent—could either not be found or refused to participate in the interviews.

We observe data on 168 participants, while the potential control group of non-participants consists of 195 individuals. After dropping records with missing values on key characteristics (e.g., employment history, unemployment duration) we are left with a sample of 288 individuals.[51] Table 4.3 illustrates the distribution of the retained sample regarding participation in either of the two program stages. Among the 142 participants, about one in three was exposed to the *Beautiful Serbia* program only through the vocational training stage, whereas about one in five was exposed to the program only through the temporary employment stage. The ratio of participants to non-participants in our working sample is close to one.

In Table 4.4 we present some descriptive statistics of the individuals subject to one of the three possible treatments—participation in the complete program, participation in the vocational training stage only, and participation in the temporary employment stage only—and of the individuals not participating at all who are potential controls. Substantial differences between participants and non-participants arise in our sample. In particular, across all treatments participants appear to be signif-

[50]Kluve et al. (1999) demonstrate that pre-unemployment labor market careers are also extremely important when assessing active labor market policies in transition economies, where variations in these histories tend to be smaller than in Western economies because they start from a situation of no formal unemployment.

[51]We also drop 7 individuals in the potential control group who exit the labor market by turning into pensioners or students.

Table 4.4: Descriptive statistics (selected variables)

| | CP | VT | TE | NP |
|---|---|---|---|---|
| Age | 31.09 | 31.85 | 33.36 | 34.23 |
| | ( 9.84) | (10.20) | (10.60) | (11.79) |
| Married | 0.3182 | 0.5000 | 0.6786 | 0.5822 |
| | (0.4693) | (0.5053) | (0.4756) | (0.4949) |
| Roma | 0.1061 | 0.2083 | 0.2143 | 0.0822 |
| | (0.3103) | (0.4104) | (0.4179) | (0.2756) |
| Belgrade | 0.4848 | 0.5000 | 0.4285 | 0.3151 |
| | (0.5036) | (0.5053) | (0.5040) | (0.4661) |
| Education: primary school or less | 0.3182 | 0.4167 | 0.3571 | 0.2877 |
| | (0.4693) | (0.4982) | (0.4880) | (0.4542) |
| Education: vocational school (3 years) | 0.3333 | 0.3333 | 0.3571 | 0.4110 |
| | (0.4750) | (0.4764) | (0.4880) | (0.4937) |
| Previous unemployment duration | 31.33 | 36.83 | 42.68 | 60.05 |
| (in months) | (37.67) | (41.78) | (50.07) | (54.69) |
| Employed at all in last 3 years | 0.7424 | 0.7292 | 0.8214 | 0.5685 |
| | (0.4407) | (0.4491) | (0.3900) | (0.4970) |
| Actively searching for a job | 0.8485 | 0.8125 | 0.8571 | 0.6370 |
| | (0.3613) | (0.3944) | (0.3563) | (0.4825) |
| # observations | 66 | 48 | 28 | 146 |

*Note:* Mean values of selected variables (standard deviation in brackets). CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

icantly younger than non-participants, better educated, more likely to belong to the ethnic group of Roma, and more likely to live in Belgrade. Furthermore, in January 2004 when the program started, the participants had experienced shorter spells of unemployment, had more frequently been employed in the past 36 months, and more often actively searching for a job than non-participants.

The substantial differences in observed characteristics indicate that the pre-matching routine to create a control group has not worked satisfactorily. One potential explanation would be that the selection of the control group was based on planned rather than on accomplished interviews. An alternative—and probably more relevant—explanation would be that the probabilities to participate in the program were actually affected by individual characteristics beyond those few used by the pre-matching routine, see above. In any case, the observed characteristics of the program participants altogether appear to give them a comparative advantage over the non-participants—in particular concerning potential labor market success. Therefore, one would expect that a comparison of mean outcomes between the two groups overestimates the positive program effects. In order to avoid this bias, we need to rely on econometric techniques for constructing a control group that is effectively comparable to the treatment group.

## 4.4 Evaluation Approach

Ideally, we would like to compare the outcomes for the individuals participating in the *Beautiful Serbia* program ($Y^1$) with the outcomes for the same individuals if they had not participated ($Y^0$). If $D$ denotes participation, where $D = 1$ if a person participates in the program and $D = 0$ otherwise, the actual outcome for individual $i$ can be written as:

$$Y_i = Y_i^1 \cdot D_i + Y_i^0 \cdot (1 - D_i) \ . \tag{4.1}$$

The individual treatment effect would be given by the difference $\Delta_i = Y_i^1 - Y_i^0$. However, it is impossible to calculate this difference because one of the outcomes is unobservable. Instead, the evaluation literature concentrates on population average gains from treatment—usually on the average treatment effect on the treated (ATT or $\Delta_{ATT}$) which is formally given by:

$$\Delta_{ATT} = E(\Delta|D = 1) = E(Y^1|D = 1) - E(Y^0|D = 1) \ . \tag{4.2}$$

It is the principle task of any evaluation study to find a credible estimate for the second term on the right hand side of equation (4.2), which is unobservable.

As metioned above, one possible solution could be to simply compare the mean outcomes of participants and non-participants. But if $E(Y^0|D = 1) \neq E(Y^0|D = 0)$, estimating the ATT by the difference between the sub-population means of these two groups will yield a selection bias—which is likely the case in our data. On the other hand, if treatment assignment is *strongly ignorable*, i.e., if selection is on observable characteristics $X$ (unconfoundedness or conditional independence assumption), and if observable characteristics of participants and non-participants overlap (common support), the matching estimator is an appealing choice to estimate the desired counterfactual (Rosenbaum and Rubin, 1983). Under these conditions, the distribution of the counterfactual outcome $Y^0$ for the participants is the same as the observed distribution of $Y^0$ for the comparison group *conditional on the vector of covariates* $X$. Formally,

$$E(Y^0|X, D = 1) = E(Y^0|X, D = 0). \tag{4.3}$$

Entering this relation into (4.2) allows estimating the ATT. Rosenbaum and Rubin (1983) show that if treatment assignment is strongly ignorable *given* $X$, it is also strongly ignorable *given any balancing score that is a function of* $X$.[52] One possible balancing score is the propensity score $P(X)$, i.e., the probability of participating in a given program.

There are several propensity score matching matching methods suggested in the literature.[53] Based on the characteristics of our data, we opt to apply nearest-neighbor matching with replacement. This matching method has the advantage of being the most straightforward matching estimator: a given participant is matched with a non-participant who is closest in terms of the estimated propensity score. As the participants and non-participants in our sample appear quite different, we allow matching with replacement to avoid bad matches between high-score participants and low-score non-participants. The disadvantage of this approach is that the variance of the estimator increases as the constructed counterfactual outcome is based on less distinct non-participants (Smith and Todd, 2005a).

For the variance of the estimated ATT, we apply the approximation suggested by Lechner (2001, 2002). The following formula applies for nearest neighbor matching with replacement:

$$Var(\hat{\Delta}_{ATT}) = \frac{1}{N_1} \cdot Var(Y^1|D = 1) + \frac{(\sum_{j \in \{D=0\}} (w_j)^2)}{(N_1)^2} \cdot Var(Y^0|D = 0) \ , \tag{4.4}$$

where $N_1$ is the number of matched treated individuals and $w_j$ is the number of times individual $j$ from the control group is used. We checked the accuracy of this

[52]When there are many covariates, it is impractical to match directly on covariates because of the curse of dimensionality. See, e.g., Zhao (2008) for some comments on this problem.

[53]See, e.g., Caliendo and Kopeinig (2008) for an overview.

approximation by also calculating the variance of the estimated treatment effects based on bootstrapping procedures. Although nearest neighbor matching does not satisfy the basic conditions for the bootstrap and the bootstrap variance diverges from the actual variance (Abadie and Imbens, 2008), this alternative method gives similar variances of the estimated treatment effects and does not change the implications presented below.

We estimate the probability of treatment in the *Beautiful Serbia* program conditional on observable characteristics—the propensity score—using binary probit models with participation as the dependent variable.[54] The potential control group always consists of the individuals who did not participate in the program at all. Our preferred specifications of the propensity score include a full range of personal characteristics. We measure regional variation in program participation rates by including an index variable taking the value of one if an individual lives in Belgrade and zero otherwise. However, as all participants in Belgrade entered the program in 2004, and almost all participants outside Belgrade entered in 2005, this variable also captures the variation related to the timing of program entry.[55]

Table 4.5 depicts the marginal effects of the probit estimates underlying the propensity scores for the three treatments. The results basically confirm the impression from the descriptive statistics. It appears that individuals relatively close to the labor market—i.e., individuals of younger age, relatively short-term unemployed, recently employed or actively searching for a job—had a higher chance to participate in the *Beautiful Serbia* program.

The distributions of the propensity scores obtained from the probit estimates are on display in Figure 4.1. A comparison of participants and non-participants reveals that the latter tend to be endowed with characteristics that make them systematically less likely to be selected for participation in the *Beautiful Serbia* program. Among the individuals participating in both stages of the program, 4 have a higher propensity score than the individual with the highest estimated propensity score among the non-

[54]Estimations are done using the `PSMATCH2` Stata ado-package by Leuven and Sianesi (2003).

[55]We have tried several specifications of the probit model. The results did not change qualitatively. For instance, including the number of (small) children living in the household does not change the predictions since all individuals in our sample are men for whom age and marital status already capture most of the effect possibly associated with children. Our preferred specifications appear to deliver the best overall predictions of program participation rates.

Table 4.5: Probit estimates (marginal effects)

| | CP vs. NP | VT vs. NP | TE vs. NP |
|---|---|---|---|
| ln(Age) | -98.8535 * | -40.0550 | 7.4593 |
| ln(Age)$^2$ | 29.3678 * | 12.2453 | -1.7225 |
| ln(Age)$^3$ | -2.8845 * | -1.2343 | 0.1222 |
| Married | -0.2717 *** | -0.1577 * | 0.0331 |
| Roma | 0.1607 | 0.2182 | 0.1851 * |
| Belgrade | 0.0538 | 0.1339 * | 0.0658 |
| Homeowner | 0.1540 | 0.1931 ** | 0.4381 *** |
| Education: primary school or less | 0.1157 | 0.1577 * | 0.0442 |
| Education: vocational school (3 years) | 0.0127 | 0.1022 | 0.0534 |
| Disabled | | -0.0254 | -0.0046 |
| Mobile | -0.1169 | -0.1675 ** | -0.0634 |
| Unemployed $\leq$ 12 months | 0.4141 *** | 0.2587 ** | 0.2331 *** |
| Unemployed 13–24 months | 0.2749 *** | 0.2156 *** | 0.0178 |
| Unemployed 25–36 months | 0.2165 ** | 0.4246 ** | 0.1222 |
| Unemployed 37–48 months | 0.2712 ** | 0.1826 | 0.1062 |
| Employed in last 3 years | 0.0959 | 0.0975 | 0.2091 *** |
| Share of employment in last 3 years | -0.1704 | -0.2355 * | -0.2510 ** |
| Other income | -0.2669 | -0.2005 | -0.1196 |
| Jobsearcher | 0.1971 ** | 0.1296 * | 0.1061 ** |
| ALMP | -0.2503 ** | -0.1240 | |
| Jobdesire | 0.1386 | 0.0597 | 0.0612 |
| Jobchances | 0.0594 | 0.2020 ** | -0.0479 |
| Jobchances × Jobdesire | 0.0551 | | |
| Jobchances × Employed in last 3 years | -0.0381 | | |
| Jobchances × Roma | -0.0246 | | |
| Roma × Belgrade | 0.5449 | 0.0419 | |
| Roma × Homeowner | -0.2600 | -0.1661 ** | |
| Roma × Married | | 0.0991 | |
| Mobile × Education: primary school or less | -0.2494 | | |
| Jobsearcher × Unemployed 25–36 months | | -0.1029 | |
| Employed in last 3 years × Homeowner | | | -0.1156 ** |

*Note:* *** significant at 1%; ** significant at 5%; * significant at 10%. CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

participants. Hence these individuals are off support according to the usual 'Minmax' criterion and need to be excluded for the computation of the ATT. To achieve common support, we need to exclude 5 (3) observations when evaluating participation in the vocational training (temporary employment) stage only.

After forming the matched pairs, a suitable way to assess the matching quality is comparison of the standardized bias before matching, $SB^b$, to the standardized bias after matching, $SB^a$. The standardized biases are defined as

$$SB^b = \frac{(\overline{X}_1 - \overline{X}_0)}{\sqrt{0.5 \cdot (V_1(X) + V_0(X))}} \quad ; \quad SB^a = \frac{(\overline{X}_{1M} - \overline{X}_{0M})}{\sqrt{0.5 \cdot (V_{1M}(X) + V_{0M}(X))}} \ , \qquad (4.5)$$

where $X_1$ ($V_1$) is the mean (variance) in the treated group before matching and $X_0$ ($V_0$) the analogue for the comparison group. $X_{1M}$ ($V_{1M}$) and $X_{0M}$ ($V_{0M}$) are the corresponding values after matching (Rosenbaum and Rubin, 1985). Following the example of Sianesi (2004) we also re-estimate the propensity score on the matched sample to compute the pseudo-$R^2$ before and after matching.

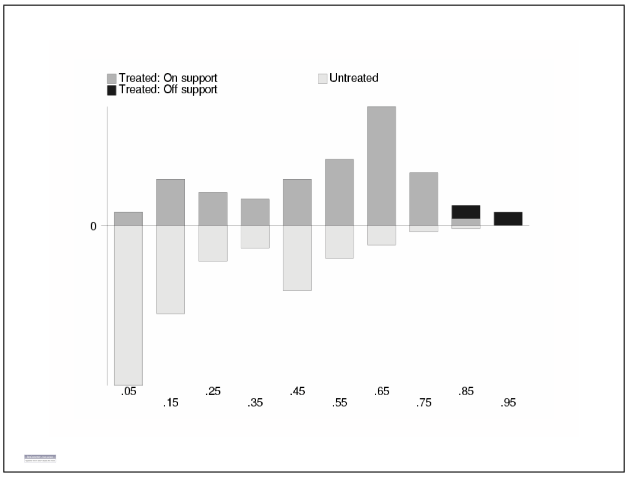
These measures (see Table 4.6) suggest that the quality of our matching procedures is quite satisfactory. The standardized bias of the matched sample is markedly smaller than that of the unmatched sample. Likewise, the pseudo-$R^2$ after matching are fairly low and decrease substantially compared to before matching. This is what we should expect considering that after matching, there should not be any systematic differences in the distribution of observable characteristics between participants and matched non-participants. Therefore, the test of the matching quality makes us confident to estimate meaningful treatment effects on the basis of nearest neighbor matching with replacement, despite of the rather small sample available for constructing the matched pairs.

## 4.5 Treatment Effects

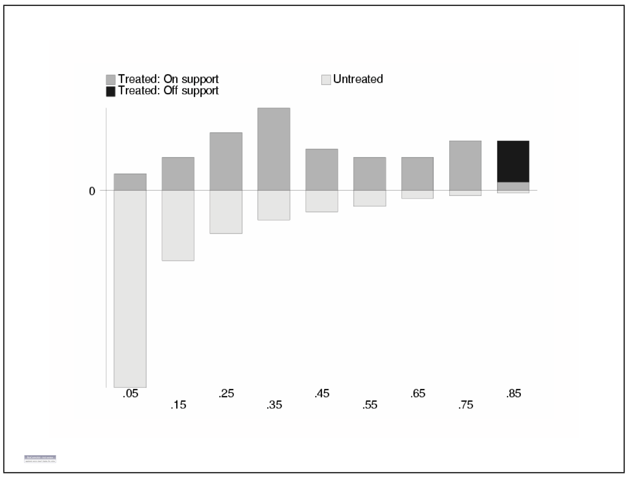
In the following we first adopt the conventional perspective of ALMP evaluation studies and assess the causal impact of the *Beautiful Serbia* program on labor market outcomes, i.e., on unemployment and employment probabilities. In a second step, we will look at subjective well-being variables at the core of our interest.

Figure 4.1: Distribution of propensity scores, common support

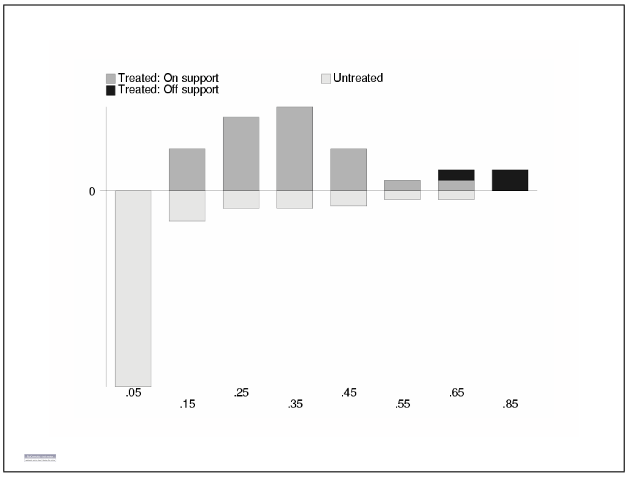
(a) CP vs. NP




(b) VT vs. NP




(c) TE vs. NP




*Note:* CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

Table 4.6: Matching quality

| | CP vs. NP | VT vs. NP | TE vs. NP |
|---|---|---|---|
| # treated individuals | 66 | 48 | 28 |
| # treated individuals off support | 4 | 5 | 3 |
| # matched pairs | 62 | 43 | 25 |
| Mean SB before matching | 0.1962 | 0.2467 | 0.1965 |
| Mean SB after matching | 0.0862 | 0.0882 | 0.1001 |
| Pseudo-$R^2$ before matching | 0.2573 | 0.2872 | 0.2890 |
| Pseudo-$R^2$ after matching | 0.1363 | 0.1688 | 0.1139 |

*Note:* The mean SB is calculated as the mean of the single characteristics' SB (in percent). CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

Table 4.7: ATT labor market outcomes

| | CP vs. NP | VT vs. NP | TE vs. NP |
|---|---|---|---|
| Unemployment | -0.1290 | -0.0698 | 0.0000 |
| Regular job | 0.1290 | 0.0465 | 0.1200 |
| Seasonal job | -0.0161 | 0.0930 | -0.1600 |
| ALMP job | 0.0323 | -0.0698 | 0.0400 |
| # matched pairs | 62 | 43 | 25 |

*Note:* *** significant at 1%; ** significant at 5%; * significant at 10%. CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

### 4.5.1 Labor Market Outcomes

The survey data do not trace individuals' employment histories. Hence our outcome variables are based on the labor market status at the time of the interview. More precisely, we consider four different labor market states: (a) unemployment, (b) employment in a regular job including self-employment, (c) employment in a seasonal job, and (d) employment in another active labor market program implemented by the National Employment Service ('ALMP job'). Table 4.7 summarizes the estimated ATT for these four different labor market outcomes and the three possible treatments.

Our point estimates suggest that participation in both stages of the program reduces the probability of being unemployed at the survey date by about 13 percentage points, compared to not participating in the program. Participation in the training stage only reduces the unemployment rate by 7 percentage points, whereas participation in the temporary employment stage apparently has no effect on the propensity of being unemployed. The latter result is estimated on very few observations, however. In fact, none of the estimated ATT is statistically significant at conventional levels. In general, the small scale of the program and therefore small sample sizes will only yield significant ATT if participants and matched non-participants exhibit very distinct outcomes.

Considering overall employment, the ATT basically mirror those concerning unemployment. Participation in the program is generally associated with a higher employment rate. However, some differences appear between the different treatments concerning the type of employment.[56] Participation in the complete program positively affects the chances of working in a regular job while the chances of seasonal employment or of an ALMP job basically remain unaffected. In contrast, for participation in the training stage of the program only, the strongest program impact is on employment in a seasonal job. The effect on seasonal employment is even stronger than the overall employment effect: Participation in the training measure reduces the chance to become employed in another active labor market program. Finally, while participation in the temporary employment stage of the program only has basically no effect on the overall employment rate, it seems to affect the type of employment: Treated individuals appear to work more frequently in regular jobs and less frequently in seasonal jobs.

In sum, our findings concerning the impacts of the Beautiful Serbia program on labor market outcomes suggest that both the vocational training and the temporary employment part (and therefore the program taken as a whole) exert a positive influence on the employment prospects of the participants. However, the positive effects are not sufficiently strong or clear-cut to be considered statistically significant.

[56] The overall ATT concerning employment is the sum of the ATT regarding the three different types of employment.

### 4.5.2 Subjective Well-Being

Even if an active labor market program does not immediately raise the employment probabilities of participants, a social planner may find it beneficial if it manages to improve the individual welfare of the target group. The survey data collected in connection with the *Beautiful Serbia* program provide us with the unique opportunity to study program impacts also on various dimensions of life that may serve to approximate individual well-being or 'happiness'.

In the literature, happiness is usually measured by the answer to a very broad question. For instance, the U.S. General Social Survey asks: *Taking things all together, how would you say you are these days—would you say you are very happy, pretty happy, or not too happy?* The individuals in our data were not asked for such a global assessment of their whole sphere of life. Instead, we observe answers to a set of questions relating to items that give a reasonable picture of how their personal situation concerning various aspects of life has changed over time. Individuals were requested to compare their situation at the time of the interview with that before the *Beautiful Serbia* program came into effect, and had to judge whether their situation has strongly or somewhat improved, has stayed more or less the same, or has strongly or somewhat deteriorated.[57] In detail, the survey requested a self-assessment of changes concerning self-confidence, the desire to find a job, social contacts, health status, the family income situation, personal qualification and skills, and the chances to find a regular job. These items have been identified as determinants of personal happiness (Frey and Stutzer, 2002). However, the extent to which the different items are related to subjective well-being varies. For example, personal health ratings and happiness appear to be highly correlated, whereas changes in income are considered to have only temporary impacts on subjective well-being, probably due to the phenomenon of adaptation (Layard, 2006). The dimensions of 'qualification and skills' and 'job chances' contain information on *employability*, which is a more general concept than

[57]Individuals from Belgrade where the program was introduced earlier were asked to compare their situation to that in the beginning of 2004, and individuals from in Niš and Zrenjanin where the program started later were asked to compare their situation to that in the beginning of 2005.

Table 4.8: ATT indicators of subjective well-being.

| | CP vs. NP | | VT vs. NP | | TE vs. NP | |
|---|---|---|---|---|---|---|
| Self-confidence | 0.1129 | | 0.2093 | * | 0.2800 | ** |
| Job desire | 0.2419 | ** | 0.2558 | ** | 0.1200 | |
| Social contacts | 0.1451 | | 0.1860 | | 0.2800 | ** |
| Qualification and skills | 0.3387 | *** | 0.5116 | *** | 0.2400 | ** |
| Health | 0.1774 | ** | 0.0233 | | 0.0000 | |
| Job chances | 0.1129 | | 0.0698 | | 0.2400 | *** |
| # matched pairs | 62 | | 43 | | 25 | |

*Note:* *** significant at 1%; ** significant at 5%; * significant at 10%. CP indicates participation in both the vocational training and the temporary employment stage of the program, VT (TE) indicates participation in the vocational training (temporary employment) stage of the program only, NP indicates non-participation.

actual employment.[58] An improvement of subjective employability probably reduces the psychic cost of being unemployed, and thus may put individuals higher on the happiness scale.

In our subsequent analysis, we apply our matching approach to the subjective data. As outcome variables, we define dummy variables that take the value of one if individuals report that their personal situation has strongly or somewhat improved, and take a value of zero otherwise. In this way, the ATT measure the change in the percentage share of individuals judging their personal as improved because of program participation. Table 4.8 summarizes our findings.

The general impression based on the point estimates is that program participation has improved the personal situation with regard to all aspects of life considered. In contrast to the impact on labor market outcomes, the program effects often appear so substantial that the estimated ATT are statistically significant despite the small sample sizes on which they are estimated. For all treatments, the strongest program impact is on the subjective rating of qualification and skills. The impact is particularly strong for participants in the vocational training stage only, followed by participants in both stages of the program. This means that the vocational training content of the program is viewed positively from the participants' perspective even when it does not

[58] Following McQuaid and Lindsay (2005), the concept of employability we have in mind is a narrow one—thereby focusing on individual factors, i.e., essential attributes and personal competencies—and close to the operational versions of the 'socio-medical employability' and the 'manpower policy employability'.

immediately raise the employment rate. Among those individuals who participated in both stages of the *Beautiful Serbia* program, the share with improved job desire and improved self-assessed health is significantly higher than among comparable individuals who were not affected by the program. Similar positive effects arise considering those participating only in the training stage, which furthermore appears to significantly improve self-confidence. A strong self-confidence effect also occurs for those participating in the temporary employment stage only. Personal relations established at the work or training place are probably responsible for the clear growth of social contacts (15–28 percentage points) achieved through the *Beautiful Serbia* program.

It is interesting to note that the program—though focused on the construction sector offering probably relatively poor working conditions—if anything positively impacts on subjective health status. Among the individuals participating in the whole program, the rate of those reporting an improvement in health compared to the pre-program situation is 18 percentage points larger than among non-participants, and the effect is statistically significant. The ATT concerning health status are much smaller for the other two treatments, but, judged by the point estimates, they are at least not negative.

Taken together, the positive program effects considering individuals' subjective assessment of conditions of life appear to be larger than the program impacts when considering their objective labor market status. This suggests that the program improves subjective well-being through other channels than the labor market. The impacts we find are strong for all treatments considered. Even the subjective health rating—a key determinant of happiness—significantly increases for those going through the complete program.

## 4.6 Conclusion

This chapter evaluates the *Beautiful Serbia* program providing vocational training and temporary employment to disadvantaged unemployed. While using standard matching techniques to bring out causal average treatment effects on the treated, the analysis deviates from routine program evaluation by considering subjective measures of individual well-being as possible outcomes. Hence this chapter is linked to the rising economic literature focused on the concept of happiness as an approximation for the individual welfare scale.

Given that the ultimate goal of social polices is the improvement of individual welfare, subjective well-being clearly is a relevant dimension for a full impact assessment of an active labor market program. The evaluation results obtained from the *Beautiful Serbia* program indeed provide an example that the positive effects of a policy can appear stronger if it is judged by subjective well-being rather than by labor market effects. The program probably impacted on individual welfare through other channels than the immediate economic status, notably by strengthening self-confidence, job desire and social inclusion of the participants.

Unfortunately, due to the small scale of the program and certain deficiencies in the accomplished survey, the treatment effects estimated for the *Beautiful Serbia* program overall allow only tentative conclusions. The systematic inclusion of subjective measures of well-being into the evaluation of a larger-scale program as well as the inclusion of more direct measures of the individual happiness scale that are also tested for behavioral relevance thus remain a challenge for future research.

# Chapter 5

# You Live and Learn: Private-Sector Training in Germany

*Nowadays the need—and the possibility—to extend the working life becomes more and more apparent in many industrialized countries. This chapter focuses on the specific situation in Germany, where for many decades the dominant view had been that school education and initial training provide a sufficient endowment for the entire working life. Further qualifications, if necessary at all, could be obtained through experience on-the-job. However, this view has changed. Lifelong learning is now viewed as a necessary complement to school education and initial training in order to acquire and to update vocational skills and qualifications. We investigate the incidence, wage effects and employment effects of private-sector training in Germany for two different periods: (a) from 1997 to 2000 and (b) from 2001 to 2004.*

## 5.1 Introduction

Nowadays the speed with which knowledge changes is enormous. In addition, the need—and the possibility—to extend the working life becomes more and more apparent. Both of these facts necessitate a broader provision of training as, e.g., Brenke and Zimmermann (2005) point out. Although this consequence applies to many industrialized countries, this chapter focuses on the specific situation in Germany.

We concentrate on training activities of employed individuals, and more specifically on their effectiveness. This chapter thus complements the previous chapters in which training activities of unemployed individuals are studied. Training for employed individuals is frequently referred to as private-sector or workplace training. In the literature the distinction between on-the-job and off-the-job training is also often found in this context, but its definition varies and thus does not seem helpful for the purpose of this chapter. While dictionaries typically define on-the-job training as being located at the workplace and supervised by experienced workers (Rutherford, 1992; Black, 2002), some studies define on-the-job training as financed or arranged by the employer (e.g., Evertsson, 2004) and other studies regard the formality or informality of training as the distinctive feature (e.g., Dearden et al., 2000). In the course of this chapter we thus avoid the terms on-the-job and off-the-job training and use private-sector or workplace training instead.[59]

For many decades the dominant view in Germany had been that school education and initial training provide a sufficient endowment for the entire working life. Further qualifications, if necessary at all, could be obtained through experience on-the-job. However, this view has changed. Lifelong learning is now viewed as a necessary complement to school education and initial training in order to acquire and to update vocational skills and qualifications. This notion has also become an integral part of the federal government's and federal states' agenda. For instance, in 2001 two programs were launched by the Federal Ministry of Education and Research (BMBF) to provide a more systematic approach for stimulating lifelong learning.

[59]Leuven (2005) uses private-sector and on-the-job training interchangeably. According to his definition, both activities exclude formal education, training of the unemployed and learning activities that workers undertake independently from their employer.

While the provision of (public) training for unemployed individuals and for those at risk of unemployment is well developed in Germany, the training sector for employed individuals is not well developed in Germany compared to other European countries.[60] According to Figure 5.1, the participation rate of employed individuals in non-formal education and training is comparatively low in Germany. It amounts to only 16 percent in 2003 which is clearly below the average of 25 member states of the European Union (EU25). In this context, non-formal education and training includes learning activities which are not part of a formal education program.[61] On the other hand, the conditional average intensity of these training activities—given that an employed individual participates in non-formal education—is 74 hours in Germany and thus exceeds the EU25 average. Figure 5.1 moreover indicates that Sweden, Finland and Denmark exhibit the highest participation rates across the EU25, while Hungary exhibits the highest (conditional) intensity.

Figure 5.2 displays the development of participation rates of employees in training activities between 1979 and 2003 in Germany.[62] In West Germany, for which data are available over the whole period, the share of participants has been almost steadily increasing during the 1980s and 1990s. Starting in 1997, participation rates decreased slightly. A similar pattern can be observed in East Germany, for which information is however only available from 1991 onwards: Participation rates increased during the 1990s and started to decrease after 1997. While the share of participants has been consistently higher in East Germany with a maximum of 37 percent in 1997, participation rates in East and West Germany were exactly the same in 2003 (26 percent). This latter rate corresponds to about 13 million individuals who participated in workplace training in that year (Kuwan and Thebis, 2004).

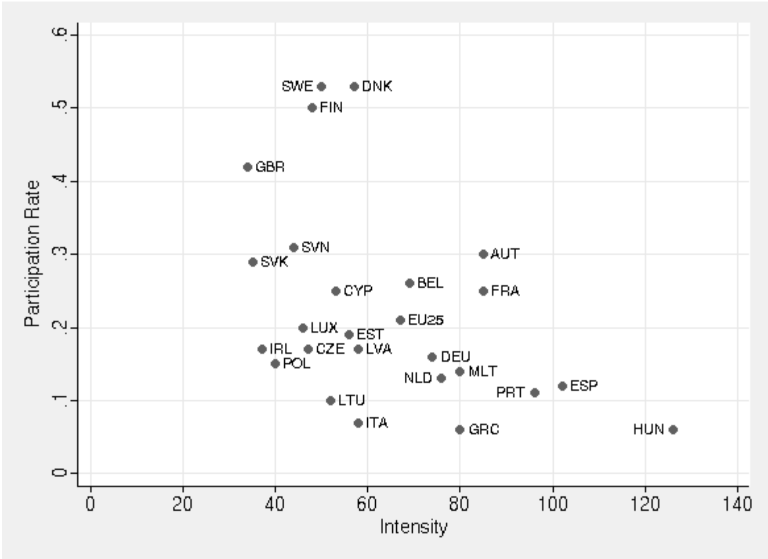
The theoretical literature on private-sector training is large. It is comprehensively surveyed by Leuven (2005). The first milestone was set when Jacob Mincer

[60] Public training for the unemployed constitutes the most important part of Germany's active labor market policy (Eichhorst and Zimmermann, 2007). Expenditures on this type of training are also high in comparison to other member states of the European Union (Melis, 2006). For recent studies on its effectiveness see, e.g., Biewen et al. (2007), Lechner and Wunsch (2008), and Rinne et al. (2007).

[61] In contrast, formal education and training corresponds to learning activities which are part of the regular system of schools, universities and colleges.

[62] This definition of training includes both general and specific training provided in courses, seminars or lectures.

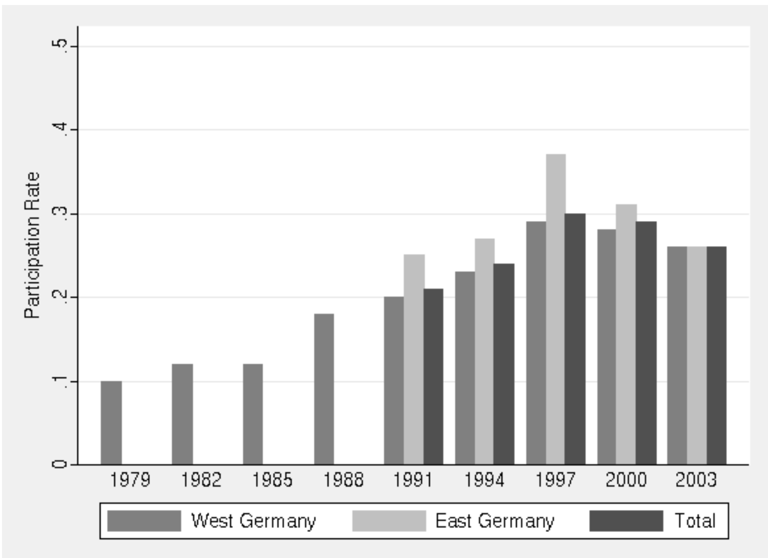
Figure 5.1: Participation Rates and Intensity in the European Union (2003)




*Note:* Participation Rates (in percent) and mean intensity (in hours taught) of non-formal education and training in 25 member states of the European Union in 2003. Only employed individuals aged 25–64 years are considered. Non-formal education and training includes learning activities which are *not* part of the regular system of schools, universities and colleges.
*Abbreviations:* EU25: EU average (25 member states); BEL: Belgium; CZE: Czech Republic; DNK: Denmark; DEU: Germany; EST: Estonia; GRC: Greece; ESP: Spain; FRA: France; ITA: Italy; IRL: Ireland; CYP: Cyprus; LVA: Latvia; LTU: Lithuania; LUX: Luxembourg; HUN: Hungary; MLT: Malta; NLD: Netherlands; AUT: Austria; POL: Poland; PRT: Portugal; SVN: Slovenia; SVK: Slovakia; FIN: Finland; SWE: Sweden; GBR: United Kingdom.

*Source:* Kailis and Pilos (2005).

Figure 5.2: Participation Rates in Germany (1979–2003)




*Note:* Participation rates (in percent) in workplace training which includes both general and specific training provided in courses, seminars or lectures. Data for East Germany is only available from 1991 onwards.

*Source:* Kuwan and Thebis (2004), p. 22.

highlighted its importance for a worker's human capital stock and estimated that this type of training accounts for at least half of it (Mincer, 1962).[63] About the same time, the formalization of the human capital theory started with Becker (1962, 1993). This literature can still be considered as the dominant perspective on private-sector training. Importantly, the distinction between general training and specific training has been established since then. Its two main predictions can be summarized as follows: *a*) under perfect competition workers receive all the returns to general training and also pay for this type of training (directly or through lower wages), while liquidity constraints can lead to under-investment; and *b*) firms finance specific training and the returns of this type of a training might be shared between the firms and workers to reduce inefficient turnover.

Motivated—among other things—by the empirical observation that there are many instances in which firms bear significant fractions of the costs of general training, the assumption of perfect competition was subsequently dropped and non-competitive theories of training evolved. These theories focus on market imperfections and information asymmetries and are summarized by Acemoglu and Pischke (1999). Major results of this line of research are the following: *a*) liquidity constraints are not sufficient to ensure firm-sponsored training in the standard perfectly competitive set-up; *b*) underinvestment in training is likely to happen; and *c*) wage compression may encourage firms to pay for training.

These theoretical predictions have been analyzed in a number of empirical studies. Pfeiffer (2001) and Frazis and Loewenstein (2005) summarize the findings for Europe and the U.S., respectively. Typically, the estimates of the wage returns to private-sector training are found to be very high. For Germany, there are already a number of related empirical studies using the same data source used in this chapter including Christensen (2001), Pischke (2001), Büchel and Pannenberg (2004) and Sauermann (2006). But these studies either focus on earlier periods (the former three ones) or focus on a specific aspect of the incidence of workplace training (the latter one which explicitly accounts for fixed-term contracts).

This chapter studies the incidence and effects of private-sector training on wages

[63]Previous contributions include Pigou (1912) and Rosenstein-Rodan (1943).

and employment stability for two different periods of time: *a*) between 1997 and 2000 and *b*) between 2001 and 2004. Next to investigating the incidence of training in cross-sectional regressions, we exploit the longitudinal dimension of our data. We estimate fixed effect models and random growth models to assess the wage effects of training as well as of different types of training. Furthermore, results on the effects of participation in private-sector training on subsequent employment are presented.

Our results indicate a fairly similar pattern with regard to the incidence of private-sector training in Germany in both periods. However, the picture which arises with respect to the effects of private-sector training on wages is relatively unstable. While we find positive wage effects of about 4–6 percent in both samples in the fixed effects specifications, these effects generally decrease quite substantially in the fixed growth rates specifications. Finally, our results indicate positive effects of participation in private-sector training on subsequent employment prospects, which seems to be solely based on whether or not an individual engaged in training at all and not on the respective duration of training.

The remainder of this chapter is organized as follows. Section 5.2 describes the data of this study. The econometric approach is presented in Section 5.3, and the results are discussed in Section 5.4. Finally, Section 5.5 concludes.[64]

## 5.2 Data

The data of this study comes from the German Socio-Economic Panel Study (SOEP).[65] The SOEP is a representative longitudinal study of private households in Germany. Its first wave started in 1984, and currently wave 24 is available which covers 2007. In that year, nearly 11,000 households were included, and more than 20,000 persons sampled.

The SOEP data cover a wide range of subjects which are included annually as well as subjects covered in modules of the survey which are not collected in every wave. The latter applies to the module "training" on which data has been collected in 1989

[64]Section 5.6 (Appendix) contains additional tables.

[65]See Wagner et al. (2007) for a comprehensive description of this data set.

(Pischke, 2001), 1993, 2000 (Wilkens and Leber, 2003; Büchel and Pannenberg, 2004) and 2004 (Sauermann, 2006). In this module, information about training activities is collected retrospectively for a period of three years prior to the interview. Individuals aged between 16 and 64 are the target population. Besides data on whether or not the individual has participated in any kind of work-related training, additional information is collected on the total number of training courses during the 3-year-period as well as more specific details about the three most recent courses. These details include start dates, duration, intensity, goals and financing. In what follows, we concentrate on the training information included in 2000 and 2004.

Focusing on that information, we can however not create a panel data set covering the period from 1997 to 2004. Since the training information is collected only for the three years prior to the respective interview date, our data does not cover the training activities which took place between 2000 and 2001. Therefore, we create two different samples: *a*) Sample A covers the period from 1997 to 2000 for which training information was retrospectively collected in 2000, and *b*) Sample B covers the period from 2001 to 2004 for which information was collected in 2004. Both samples are restricted to individuals who were included in the target population of the training questions in 2000 or 2004, respectively. Additionally, we only keep information on individuals who are between 25 and 55 years old and were employed during the entire observation period, i.e., from 1997 to 2000 or from 2001 to 2004, respectively.[66] After dropping observations with missing information in important characteristics, we end up with 2,394 individuals (9,576 observations) in Sample A and 3,432 individuals (13,728 observations) in Sample B.

Table 5.1 displays descriptive statistics of these samples at the beginning of the respective observation period. The basic socio-demographic characteristics of individuals are fairly similar in both samples. However, individuals in Sample B are about 1.5 years older, more experienced and better educated than in the sample for the earlier period.[67] The share of male and full-time employed individuals is lower in Sample B, while there are more German persons in this sample. With respect to training in-

[66] More precisely, we exclude unemployed, inactive, and self-employed individuals. Furthermore, we do not consider persons in full-time education as well as individuals working in military service, in agricultural or fishing industries, or as civil servants.

[67] Potential experience is $\max(\text{age} - \text{years of schooling} - 6, 0)$.

Table 5.1: Descriptive statistics

| | Sample A (1997–2000) | Sample B (2001–2004) |
|---|---|---|
| Male | 0.594 | 0.578 |
| | (0.491) | (0.494) |
| Age | 38.001 | 39.471 |
| | (7.431) | (7.094) |
| German | 0.870 | 0.913 |
| | (0.337) | (0.282) |
| Full time | 0.867 | 0.830 |
| | (0.339) | (0.376) |
| East Germany | 0.286 | 0.247 |
| | (0.452) | (0.431) |
| Years of schooling | 11.962 | 12.261 |
| | (2.497) | (2.471) |
| Potential Experience | 20.039 | 21.210 |
| | (7.848) | (7.646) |
| Training participant | 0.340 | 0.337 |
| | (0.474) | (0.473) |
| Number of courses | 1.282 | 1.219 |
| | (2.766) | (2.725) |
| # observations | 2,394 | 3,432 |

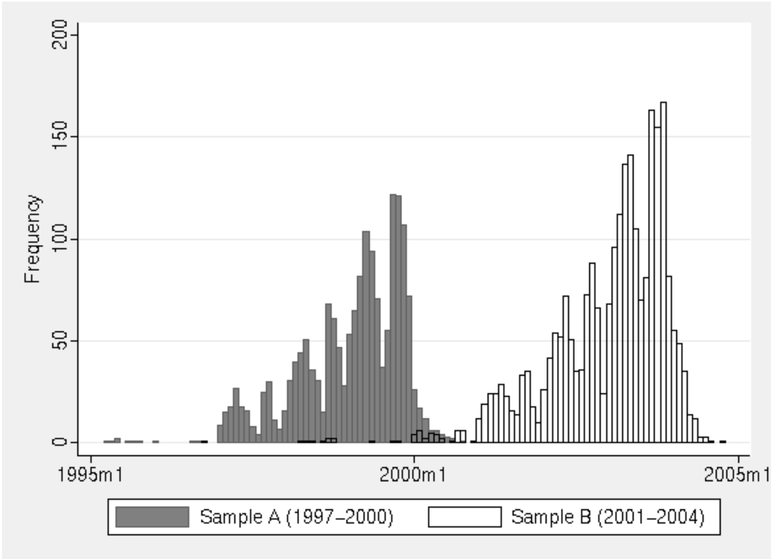
*Note:* Standard errors in brackets.

cidence the descriptive statistics are virtually the same in both samples: Roughly 34 percent have received any kind of training and an individual has participated on average in about 1.2 courses.

Although the training questions explicitly refer to the three years prior to the respective interview date, some courses are reported to have started earlier. Figure 5.3 depicts the distribution of start dates in our working samples. Courses which have started before the respective observation period, i.e., before 1997 or 2001, are only considered in what follows as far as they ended during the observation period.

Table 5.2 displays more detailed descriptive information about the incidence, duration and various types of training at the individual level. Firstly, the share of individuals participating in any kind of training between two subsequent interviews increases during the respective observation period. This may—at least in part—reflect that the training information is collected retrospectively and that information about the timing of training is only available for the three most recent courses. The descriptive information moreover reveals that most of the training is taught with less than 15 hours per week and aimed to adjust the skills and qualifications to the standards

Figure 5.3: Start dates of courses




*Note:* Start dates of the three most recent courses prior to the respective interview date.

of the current job. Furthermore, the majority of training in our data takes place during leisure hours. Concerning the financing of private-sector training activities, the descriptives show that the majority is financed by the employer alone in Sample A. In Sample B, which covers the period from 2001 to 2004, the sources of financing are relatively evenly distributed across the four categories we consider. Finally, we can distinguish between general and specific training in Sample B. Clearly, most of the training activities are categorized as specific training by the participants.

Table 5.2: Training participation

| | Sample A | | | Sample B | | |
|---|---|---|---|---|---|---|
| | 1997/1998 | 1998/1999 | 1999/2000 | 2001/2002 | 2002/2003 | 2003/2004 |
| Participation | 0.081 | 0.190 | 0.261 | 0.082 | 0.171 | 0.262 |
| | (0.272) | (0.393) | (0.439) | (0.274) | (0.377) | (0.440) |
| Duration | 3.853 | 6.999 | 9.185 | 3.267 | 5.177 | 7.091 |
| | (30.251) | (37.987) | (44.272) | (27.153) | (34.416) | (37.500) |
| Duration in full-time course | 0.326 | 0.655 | 0.712 | 0.308 | 0.567 | 0.748 |
| | (6.510) | (6.136) | (4.689) | (7.145) | (7.226) | (6.912) |
| Duration in part-time course | 0.491 | 0.749 | 1.138 | 0.388 | 0.453 | 0.787 |
| | (11.808) | (12.150) | (13.214) | (9.261) | (8.319) | (10.355) |
| Duration in course with less hours | 2.490 | 3.975 | 5.508 | 1.998 | 3.342 | 4.296 |
| | (24.638) | (28.753) | (35.899) | (21.300) | (29.780) | (30.611) |
| Duration in correspondence course | 0.414 | 1.378 | 1.778 | 0.306 | 0.417 | 0.911 |
| | (10.690) | (19.638) | (22.047) | (8.779) | (9.067) | (15.994) |
| Duration in course with aim 1 | 0.009 | 0.000 | 0.066 | 0.253 | 0.158 | 0.219 |
| | (0.429) | (0.000) | (3.068) | (8.428) | (6.870) | (8.704) |
| Duration in course with aim 2 | 0.093 | 0.222 | 0.210 | 0.006 | 0.055 | 0.149 |
| | (3.306) | (6.525) | (5.162) | (0.162) | (2.124) | (6.225) |
| Duration in course with aim 3 | 0.898 | 1.833 | 2.227 | 1.257 | 1.457 | 2.240 |
| | (14.702) | (20.734) | (23.848) | (17.955) | (19.308) | (23.071) |
| Duration in course with aim 4 | 2.402 | 4.185 | 5.685 | 1.494 | 2.954 | 3.710 |
| | (24.056) | (28.885) | (34.257) | (16.570) | (25.279) | (25.033) |
| Duration in course with aim 5 | 0.450 | 0.756 | 0.994 | 0.308 | 0.552 | 0.774 |
| | (10.774) | (12.651) | (14.696) | (9.244) | (11.677) | (12.638) |
| Duration during work hours | 1.099 | 2.615 | 3.431 | 1.217 | 1.973 | 3.584 |
| | (12.673) | (21.191) | (23.476) | (15.023) | (18.241) | (24.995) |
| Duration during leisure hours | 2.753 | 4.313 | 5.633 | 2.050 | 3.158 | 3.469 |
| | (27.563) | (31.631) | (37.693) | (22.729) | (29.284) | (28.210) |
| Duration without any financing | 0.534 | 0.837 | 1.342 | 0.185 | 0.447 | 1.052 |
| | (11.512) | (13.063) | (18.924) | (5.258) | (9.011) | (14.272) |
| Duration financed by employer | 0.703 | 1.510 | 2.695 | 0.689 | 1.297 | 2.644 |
| | (9.888) | (11.988) | (18.480) | (9.132) | (13.654) | (20.176) |
| Duration financed by employee | 1.547 | 3.139 | 4.077 | 1.381 | 2.139 | 1.988 |
| | (19.129) | (27.903) | (32.901) | (18.305) | (24.503) | (21.930) |
| Duration financed by both | 1.060 | 1.584 | 1.063 | 0.820 | 1.008 | 1.166 |
| employer and employee | (18.112) | (19.660) | (13.725) | (15.734) | (16.090) | (16.475) |
| Duration general training | | | | 0.582 | 0.748 | 1.026 |
| | | | | (10.598) | (10.833) | (9.560) |
| Duration specific training | | | | 2.686 | 4.429 | 6.030 |
| | | | | (25.058) | (32.761) | (36.323) |
| # observations | 2,394 | 2,394 | 2,394 | 3,432 | 3,432 | 3,432 |

*Note:* Standard errors in brackets. Unconditional duration of training activities (in days). Course with aim 1: re-training in a different occupation; aim 2: adjustment to standards of new job; aim 3: qualification for promotion; aim 4: adjustment to standards of current job; aim 5: other aims. Full-time courses: 30 hours per week and more; part-time courses: between 15 and 30 hours per week; less intense courses: less than 15 hours per week. Information about general and specific training only available for Sample B.

## 5.3 Econometric Approach

### Incidence

To investigate who participates in private-sector training in our data, we estimate linear probability models (LPM) with a dummy indicating participation at all in workplace training during the three years prior to the interview in 1997 or 2004, respectively. The independent variables included in these models reflect a selection of regressors which are typically included in cross-sectional earnings regressions (e.g., experience, education, firm size).

### Wage Effects

We moreover exploit the longitudinal dimension of our data. Among other things, this perspective provides a good opportunity to study the wage effects of workplace training. To assess the effects of this type of training, we follow an approach where a wage equation is augmented with training variables (Pischke, 2001).

We start by estimating standard fixed effect regressions to eliminate any effects correlated with the level of wages. Therefore, these models deal with the (potential) selection issue addressed above. Formally, the fixed effect models are as follows:

$$\ln w_{it} = X_{it}\beta + \gamma T_{it} + \alpha_i + \epsilon_{it} \ , \tag{5.1}$$

where $X_{it}$ is a set of regressors including labor market experience and tenure with the current employer, $T_{it}$ denotes the receipt of training *before* period $t$, and $\alpha_i$ is a fixed individual-specific constant affecting all time-invariant determinants of the level of wages.

However, fixed effect models may still overestimate $\gamma$. It is possible that a correlation between training and the growth rates of wages remains, in which case the fixed effect results are biased. We thus also estimate alternative models including individual specific growth rates of wages or earnings:

$$\ln w_{it} = X_{it}\beta + \gamma T_{it} + \alpha_i + \delta_i t + \epsilon_{it} \ . \tag{5.2}$$

Table 5.3: Patterns of training spells

| Sample A (1997–2000) | | | | Sample B (2001–2004) | | | |
|---|---|---|---|---|---|---|---|
| 1997/1998 | 1998/1999 | 1999/2000 | # obs. | 2001/2002 | 2002/2003 | 2003/2004 | # obs. |
| 0 | 0 | 0 | 1625 | 0 | 0 | 0 | 2333 |
| 0 | 0 | 1 | 266 | 0 | 0 | 1 | 438 |
| 0 | 1 | 0 | 65 | 0 | 1 | 0 | 90 |
| 0 | 1 | 1 | 245 | 0 | 1 | 1 | 291 |
| 1 | 0 | 0 | 30 | 1 | 0 | 0 | 44 |
| 1 | 0 | 1 | 17 | 1 | 0 | 1 | 30 |
| 1 | 1 | 0 | 50 | 1 | 1 | 0 | 65 |
| 1 | 1 | 1 | 96 | 1 | 1 | 1 | 141 |

Equations of this form are commonly referred to as random growth models (see, e.g., Heckman and Hotz, 1989). The coefficient $\delta_i$ is (roughly) the average annual growth rate—holding the explanatory variables fixed (Wooldridge, 2002).

These models are identified as long as there are at least three periods available on each individual. We already ensured this to be the case when selecting our working samples. Additionally, enough variation in the individuals' training receipt across the periods is needed since individuals who receive (the same amount of) training each year do not contribute to the estimation of $\gamma$ (Pischke, 2001). Table 5.3 displays the patterns of training spells in our working samples. Since only about 4 percent of the individuals in both samples received training between each of the waves being analyzed, the remaining participants will allow estimating the wage effects in the model with heterogenous growth rates.

## Subsequent Employment

To assess the effects of participation in private-sector training on subsequent employment, we estimate linear probability models (LPM) with a dummy indicating that the given individual is employed in the primary labor market at the respective interview date. The respective investigation period ranges from 2001 to 2007 for Sample A and from 2005 to 2007 for Sample B. Besides variables that describe the training activities in the three years prior to the investigation period, the regressions include additional control variables such as years of education, experience, unemployment rates, and GDP growth.

## 5.4 Results

This section presents the results of our empirical analyses. After analyzing the incidence of private-sector training in our samples, we present our estimates of the wage effects of this type of training.[68] Finally, the results on the effects of private-sector training on subsequent employment are displayed.

### Incidence

The first two columns of Table 5.4 present estimates of a linear probability model investigating the incidence of workplace training in the two samples. The dependent variable is a dummy variable indicating whether or not the individual receives any kind of training during the respective observation period.

Our results indicate that individuals with a German citizenship, employees of larger firms and more educated individuals are significantly more likely to receive training during both periods. The incidence of training significantly decreases with the potential experience of a given individual. After additionally controlling for the occupational level, individuals living in East Germany are significantly more likely to receive training in both samples.

### Wage Effects

Below, we will assess the effect of participation in workplace training on hourly wages (and earnings). Figure 5.4 shows the annual average gross hourly wages in our working samples. For this representation, we differentiate between individuals who receive any kind of training during the observation period and those who do not. It shows that there are substantial differences between individuals who participate in private-sector training and those who do not both in terms of the level of hourly wages as well as the growth rates. It thus seems to be appropriate taking this into account when analyzing the wage effects of private-sector training.

As mentioned above, fixed effect regressions will eliminate any effects correlated

[68] Section 5.6 (Appendix) contains estimates of the earnings effects of private-sector training.

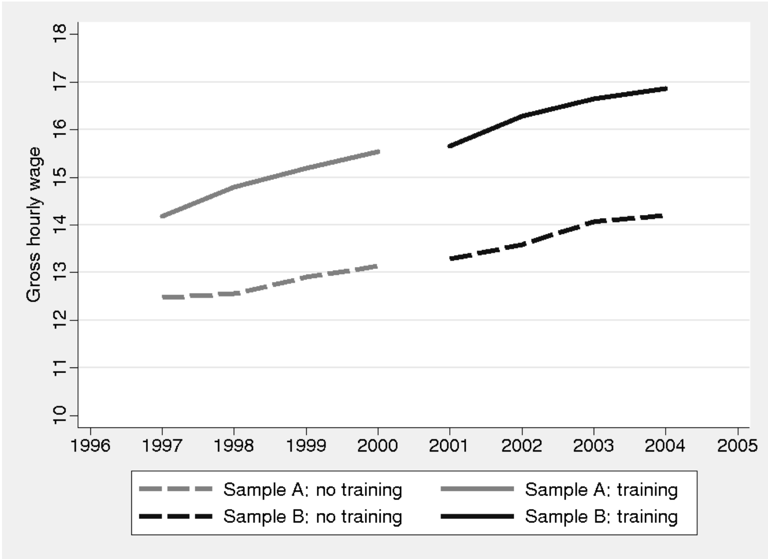
Table 5.4: Training incidence

| | Sample A (1997–2000) | | Sample B (2001–2004) | |
|---|---|---|---|---|
| | (1) | (2) | (1) | (2) |
| Years of schooling | 0.03*** | 0.007 | 0.028*** | 0.003 |
| | (0.004) | (0.005) | (0.004) | (0.004) |
| Potential experience | -.004*** | -.004*** | -.005*** | -.006*** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| Male | -.006 | 0.029 | 0.012 | 0.027 |
| | (0.022) | (0.024) | (0.019) | (0.02) |
| German | 0.175*** | 0.1*** | 0.105*** | 0.043 |
| | (0.03) | (0.03) | (0.029) | (0.029) |
| Full time | 0.05 | 0.042 | 0.032 | 0.006 |
| | (0.031) | (0.03) | (0.024) | (0.024) |
| East Germany | 0.032 | 0.06*** | 0.005 | 0.038** |
| | (0.022) | (0.022) | (0.019) | (0.019) |
| Untrained worker | | — | | — |
| | | (—) | | (—) |
| Semi-trained worker | | 0.031 | | 0.047 |
| | | (0.059) | | (0.063) |
| Trained worker | | 0.121** | | 0.137** |
| | | (0.059) | | (0.063) |
| Foreman | | 0.227*** | | 0.245*** |
| | | (0.067) | | (0.071) |
| Untrained employee | | 0.167** | | 0.148* |
| | | (0.072) | | (0.076) |
| Trained employee | | 0.263*** | | 0.197*** |
| | | (0.065) | | (0.067) |
| Qualified professional | | 0.331*** | | 0.303*** |
| | | (0.058) | | (0.063) |
| High qualified professional / managerial | | 0.356*** | | 0.38*** |
| | | (0.064) | | (0.067) |
| Firm size <20 | — | — | — | — |
| | (—) | (—) | (—) | (—) |
| Firm size 20–200 | 0.04 | 0.037 | 0.018 | 0.011 |
| | (0.027) | (0.026) | (0.022) | (0.022) |
| Firm size 200–2000 | 0.076*** | 0.066** | 0.078*** | 0.075*** |
| | (0.028) | (0.027) | (0.024) | (0.024) |
| Firm size ≥2000 | 0.156*** | 0.131*** | 0.143*** | 0.128*** |
| | (0.029) | (0.028) | (0.024) | (0.024) |
| # obs. | 2,394 | 2,394 | 3,432 | 3,432 |
| $R^2$ | 0.118 | 0.158 | 0.097 | 0.131 |

*Note:* Regressors refer to 1997 or 2001, respectively. Regressions additionally include a constant and 12 industry dummies. Standard errors in brackets.

*** significant at 1%; ** significant at 5%; * significant at 10%; — reference category.

Figure 5.4: Average gross hourly wages




*Note:* Gross hourly wages in Euros. We adjust for inflation by using the respective CPI.

with the level of wages. This type of models is therefore a way to deal with the selection issue, i.e., the selection into training based on unobservable characteristics which are related to the level of wages. Tables 5.5 and 5.6 display the results of fixed effect regressions on log hourly wages, where specifications (1)–(6) include a series of different training variables. In the first specification we estimate the annual effect of training on the hourly wage rate of about 4 percent in Sample A and 6 percent in Sample B. However, only in the latter case the estimated coefficient is significantly different from zero. The additional inclusion of a dummy variable for participating in training at all (in addition to the duration variable) slightly raises the effects in both cases. The estimated coefficient on the dummy variable is negative, which can be interpreted as evidence in favor of the notion that participation itself does not itself have a positive impact on hourly wages but the respective duration does so. Or, alternatively, very short training courses do not improve the wage prospects.

Further specifications include the respective training duration for different types of training, where we distinguish courses *a*) by intensity, *b*) by aim, *c*) whether taking place during work or leisure hours (or while being unemployed), and *d*) by sponsor. For Sample B we can moreover distinguish between general and specific training. The results of these regressions indicate that full-time training courses, i.e., courses which are taught with 30 hours per week or more, have negative wage effects in both samples—although these effects are not significantly different from zero. With respect to the aim of the courses, the picture is less clear cut. While in Sample A courses which aim to adjust skills and qualifications to the standards of the current job show significantly positive wage effects, we identify courses with other as courses with significantly positive wage effects in Sample B. With respect to the timing of the courses, the results for the two samples essentially point into the opposite directions: While training which takes place during work hours has a significantly positive wage effect in Sample A, it is training which takes place during leisure hours for which we find such an effect in Sample B. We find positive wage effects in both samples for courses which are financed by both the employer and employee, although only in Sample A the estimated coefficient is significantly different from zero. In Sample B we instead find a significantly positive wage effect of training which is financed by the employee alone. Finally, the distinction between general and specific training which we can only perform in Sample B reveals a significantly positive wage effect of specific training.

However, fixed effect models may still overestimate $\gamma$ (i.e., the effect of one year spent in training on wages) for the reasons mentioned above. Therefore, we also estimate random growth models which include individual specific growth rates of wages. Tables 5.7 and 5.8 present the results of these regressions. Compared to fixed effect models, the estimated coefficients indeed generally decrease quite substantially. While in the first specification the estimated annual effect of training on the hourly wage rate remains at about 4 percent in Sample A, it decreases and becomes virtually zero in Sample B. Further specifications which include the respective training duration for different types of training reveal the following results: *a*) in both samples, correspondence courses have positive wage effects (significantly positive only in Sample A), while all other types of courses with higher intensities have negative wage effects; *b*) courses without any financing have a significantly positive wage effect in Sample A; and *c*) general training has a positive wage effect in Sample B, but this effect is not significantly different from zero.

Table 5.5: Fixed effects log hourly wage regressions: 1997–2000

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Any training | | -.004<br>(0.01) | | | | |
| Training duration | 0.04<br>(0.03) | 0.046<br>(0.033) | | | | |
| Training duration full-time | | | -.128<br>(0.224) | | | |
| Training duration part-time | | | 0.28***<br>(0.089) | | | |
| Training duration less hours | | | 0.02<br>(0.038) | | | |
| Training duration correspondence course | | | 0.011<br>(0.064) | | | |
| Training duration aim 1 | | | | 0.869<br>(0.831) | | |
| Training duration aim 2 | | | | 0.055<br>(0.229) | | |
| Training duration aim 3 | | | | 0.0008<br>(0.056) | | |
| Training duration aim 4 | | | | 0.066*<br>(0.039) | | |
| Training duration aim 5 | | | | -.001<br>(0.087) | | |
| Training duration during work | | | | | 0.136**<br>(0.058) | |
| Training duration during leisure | | | | | 0.007<br>(0.035) | |
| Duration without any financing | | | | | | 0.028<br>(0.079) |
| Duration financed by employer | | | | | | 0.013<br>(0.085) |
| Duration financed by employee | | | | | | 0.009<br>(0.042) |
| Duration financed by both<br>employer and employee | | | | | | 0.134**<br>(0.065) |
| # obs. | 9,576 | 9,576 | 9,576 | 9,576 | 9,576 | 9,576 |
| $R^2$ | 0.016 | 0.016 | 0.017 | 0.016 | 0.016 | 0.016 |

*Note:* Unbalanced sample including 3,224 individuals. Training duration in years. All regressions also include a full set of year dummies. Standard errors in brackets.
Course with aim 1: retraining in a different occupation; aim 2: adjustment to standards of new job; aim 3: qualification for promotion; aim 4: adjustment to standards of current job; aim 5: other aims.
Full-time course: 30 hours per week and more; part-time courses: between 15 and 30 hours per week; less intense courses: less than 15 hours per week.

*** significant at 1%; ** significant at 5%; * significant at 10%.

Table 5.6: Fixed effects log hourly wage regressions: 2001–2004

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Any training | | -.015<br>(0.01) | | | | | |
| Training duration | 0.06*<br>(0.036) | 0.081**<br>(0.039) | | | | | |
| Training duration full-time | | | -.110<br>(0.187) | | | | |
| Training duration part-time | | | 0.062<br>(0.128) | | | | |
| Training duration less hours | | | 0.068<br>(0.044) | | | | |
| Training duration correspondence course | | | 0.123<br>(0.105) | | | | |
| Training duration aim 1 | | | | 0.233<br>(0.148) | | | |
| Training duration aim 2 | | | | 0.035<br>(0.345) | | | |
| Training duration aim 3 | | | | 0.025<br>(0.06) | | | |
| Training duration aim 4 | | | | 0.029<br>(0.054) | | | |
| Training duration aim 5 | | | | 0.215**<br>(0.106) | | | |
| Training duration during work | | | | | -.025<br>(0.064) | | |
| Training duration during leisure | | | | | 0.103**<br>(0.044) | | |
| Duration without any financing | | | | | | -.028<br>(0.134) | |
| Duration financed by employer | | | | | | -.016<br>(0.086) | |
| Duration financed by employee | | | | | | 0.093*<br>(0.053) | |
| Duration financed by both employer and employee | | | | | | 0.063<br>(0.075) | |
| Duration general training | | | | | | | -.045<br>(0.122) |
| Duration specific training | | | | | | | 0.07*<br>(0.038) |
| # obs. | 13,728 | 13,728 | 13,728 | 13,728 | 13,728 | 13,728 | 13,728 |
| $R^2$ | 0.008 | 0.009 | 0.009 | 0.009 | 0.009 | 0.008 | 0.008 |

*Note:* Unbalanced sample including 4,487 individuals. Training duration in years. All regressions also include a full set of year dummies. Standard errors in brackets.
Course with aim 1: retraining in a different occupation; aim 2: adjustment to standards of new job; aim 3: qualification for promotion; aim 4: adjustment to standards of current job; aim 5: other aims.
Full-time course: 30 hours per week and more; part-time courses: between 15 and 30 hours per week; less intense courses: less than 15 hours per week.

*** significant at 1%; ** significant at 5%; * significant at 10%.

Table 5.7: Fixed growth rates log hourly wage regressions: 1997–2000

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Any training | | -.085*** | | | | |
| | | (0.025) | | | | |
| Training duration | 0.041 | 0.14 | | | | |
| | (0.092) | (0.097) | | | | |
| Training duration full-time | | | -.300 | | | |
| | | | (0.467) | | | |
| Training duration part-time | | | -.298 | | | |
| | | | (0.327) | | | |
| Training duration less hours | | | -.044 | | | |
| | | | (0.117) | | | |
| Training duration correspondence course | | | 0.382** | | | |
| | | | (0.185) | | | |
| Training duration aim 1 | | | | -.499 | | |
| | | | | (1.408) | | |
| Training duration aim 2 | | | | 0.439 | | |
| | | | | (0.583) | | |
| Training duration aim 3 | | | | 0.073 | | |
| | | | | (0.184) | | |
| Training duration aim 4 | | | | -.003 | | |
| | | | | (0.115) | | |
| Training duration aim 5 | | | | 0.175 | | |
| | | | | (0.307) | | |
| Training duration during work | | | | | 0.212 | |
| | | | | | (0.174) | |
| Training duration during leisure | | | | | -.021 | |
| | | | | | (0.108) | |
| Duration without any financing | | | | | | 0.51** |
| | | | | | | (0.228) |
| Duration financed by employer | | | | | | 0.03 |
| | | | | | | (0.239) |
| Duration financed by employee | | | | | | -.068 |
| | | | | | | (0.132) |
| Duration financed by both employer and employee | | | | | | -.075 |
| | | | | | | (0.2) |
| # obs. | 7,182 | 7,182 | 7,182 | 7,182 | 7,182 | 7,182 |
| $R^2$ | 0.001 | 0.003 | 0.002 | 0.001 | 0.001 | 0.002 |

*Note:* Unbalanced sample including 3,224 individuals. The models are estimated by applying fixed effects to the differenced equation. Training duration in years. All regressions also include a full set of year dummies. Standard errors in brackets.
Course with aim 1: retraining in a different occupation; aim 2: adjustment to standards of new job; aim 3: qualification for promotion; aim 4: adjustment to standards of current job; aim 5: other aims.
Full-time course: 30 hours per week and more; part-time courses: between 15 and 30 hours per week; less intense courses: less than 15 hours per week.

*** significant at 1%; ** significant at 5%; * significant at 10%.

Table 5.8: Fixed growth rates log hourly wage regressions: 2001–2004

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Any training | | 0.013<br>(0.027) | | | | | |
| Training duration | 0.0006<br>(0.113) | -.015<br>(0.117) | | | | | |
| Training duration full-time | | | -.215<br>(0.436) | | | | |
| Training duration part-time | | | -.472<br>(0.429) | | | | |
| Training duration less hours | | | -.002<br>(0.136) | | | | |
| Training duration correspondence course | | | 0.365<br>(0.312) | | | | |
| Training duration aim 1 | | | | 0.264<br>(0.522) | | | |
| Training duration aim 2 | | | | 0.106<br>(0.861) | | | |
| Training duration aim 3 | | | | -.121<br>(0.198) | | | |
| Training duration aim 4 | | | | 0.127<br>(0.158) | | | |
| Training duration aim 5 | | | | -.323<br>(0.346) | | | |
| Training duration during work | | | | | -.073<br>(0.171) | | |
| Training duration during leisure | | | | | 0.058<br>(0.149) | | |
| Duration without any financing | | | | | | -.006<br>(0.298) | |
| Duration financed by employer | | | | | | 0.012<br>(0.22) | |
| Duration financed by employee | | | | | | 0.009<br>(0.192) | |
| Duration financed by both employer and employee | | | | | | -.006<br>(0.231) | |
| Duration general training | | | | | | | 0.094<br>(0.326) |
| Duration specific training | | | | | | | -.011<br>(0.12) |
| # obs. | 10,296 | 10,296 | 10,296 | 10,296 | 10,296 | 10,296 | 10,296 |
| $R^2$ | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 | 0.001 |

*Note:* Unbalanced sample including 4,487 individuals. The models are estimated by applying fixed effects to the differenced equation. Training duration in years. All regressions also include a full set of year dummies. Standard errors in brackets.
Course with aim 1: retraining in a different occupation; aim 2: adjustment to standards of new job; aim 3: qualification for promotion; aim 4: adjustment to standards of current job; aim 5: other aims.
Full-time course: 30 hours per week and more; part-time courses: between 15 and 30 hours per week; less intense courses: less than 15 hours per week.

*** significant at 1%; ** significant at 5%; * significant at 10%.

## Subsequent Employment

We assess the effects of participation in private-sector training on subsequent employment by estimating linear probability models with a dummy indicating that the given individual is employed in the primary labor market at the respective interview date. In these models, we include variables that describe the training activities in the three years prior to the investigation period. More specifically, we include a dummy variable indicating whether or not the individual participated in any training activities at all as well as the duration which he or she spent in training activities (in years).

The results of these regressions are shown in Tables 5.9 and 5.10. We find a positive impact of participation in private-sector training on subsequent employment probabilities in both samples. The probability of being employed in subsequent years is raised by about 2–3 percentage points. While the effect becomes significantly different from zero only in 2003 and 2004 for Sample A, i.e., 3 and 4 years after we observe the last training information, it is significantly positive immediately in Sample B. However, the positive employment effects seem to disappear after around 5 years as the results for Sample A suggest. Note that in both samples the respective duration which has been spent in training activities does not improve subsequent employment prospects. Therefore, the positive effect of private-sector training which we find seems to be solely based on whether or not an individual engaged in training at all.

Table 5.9: OLS: employed at all, Sample A

| | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 |
|---|---|---|---|---|---|---|---|
| Any Training | 0.009 | 0.018 | 0.024* | 0.03** | 0.015 | 0.008 | 0.012 |
| | (0.009) | (0.012) | (0.013) | (0.014) | (0.016) | (0.017) | (0.017) |
| Training Duration | -.003 | 0.016 | 0.03 | -.017 | -.003 | 0.018 | 0.033 |
| | (0.016) | (0.021) | (0.023) | (0.025) | (0.027) | (0.03) | (0.03) |
| # obs. | 2,394 | 2,262 | 2,189 | 2,111 | 2,028 | 1,897 | 1,819 |
| $R^2$ | 0.012 | 0.012 | 0.012 | 0.02 | 0.025 | 0.048 | 0.083 |

*Note:* Training duration in years. Regressions additionally include years of education, experience, sex, citizenship, East/West Germany, regional unemployment rate, and regional GDP growth.

*** significant at 1%; ** significant at 5%; * significant at 10%.

Table 5.10: OLS: employed at all, Sample B

| | 2005 | 2006 | 2007 |
|---|---|---|---|
| Any Training | 0.026*** | 0.014 | 0.017* |
| | (0.007) | (0.009) | (0.009) |
| Training Duration | -.019 | -.011 | 0.023 |
| | (0.015) | (0.018) | (0.018) |
| # obs. | 3,432 | 3,195 | 3,038 |
| $R^2$ | 0.011 | 0.007 | 0.015 |

*Note:* Training duration in years. Regressions additionally include years of education, experience, sex, citizenship, East/West Germany, regional unemployment rate, and regional GDP growth.

*** significant at 1%; ** significant at 5%; * significant at 10%.

## 5.5 Conclusion

This chapter analyzes the incidence, wage effects and employment effects of private-sector training in Germany. We use data from the SOEP and focus on a specific module, in which information about training activities is collected retrospectively for a period of three years prior to the interview. We concentrate on two periods: *a*) from 1997 to 2000 for which training information was retrospectively collected in 2000 and *b*) from 2001 to 2004 for which information was collected in 2004.

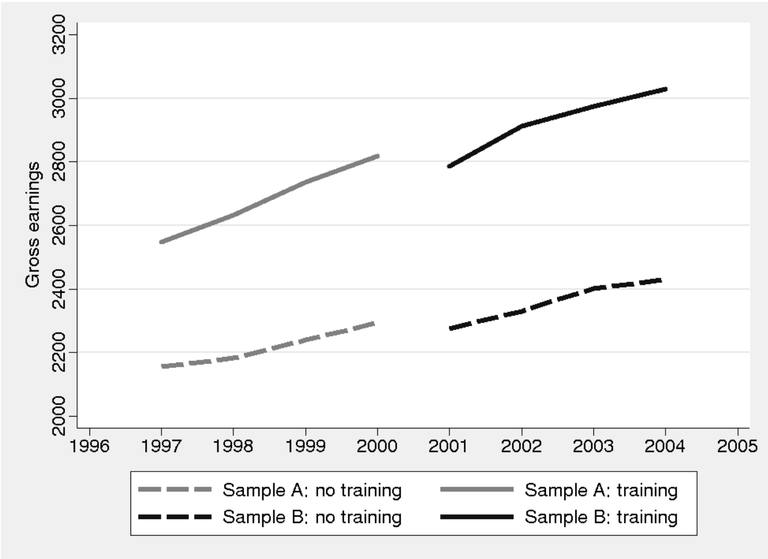
Our results indicate a fairly similar pattern with regard to the incidence of private-sector training in Germany in both periods. Individuals with a German citizenship, employees of larger firms and more educated individuals are significantly more likely to receive training during both periods. Only after additionally controlling for the occupational level, individuals living in East Germany are significantly more likely to receive training in both samples.

The picture which arises with respect to the effects of private-sector training on wages is relatively unstable. While we find positive wage effects of about 4–6 percent in both samples in the fixed effects specifications, these effects generally decrease quite substantially in the fixed growth rates specifications. More specifically, while the estimated annual effect of training on the hourly wage rate remains at about 4 percent for the earlier period, it decreases and becomes virtually zero for the later period.

With respect to the effect of participation in private-sector training on subsequent employment prospects, we find a positive effect in both samples. The probability of being employed in subsequent years is raised by about 2–3 percentage points. However, this positive employment effect seems to disappear after around 5 years. We moreover find that the respective duration which has been spent in training activities does not improve subsequent employment prospects; it thus seems that the positive employment effect of private-sector training is solely based on whether or not an individual engaged in training at all.

## 5.6 Appendix

Figure 5.5: Average gross earnings




*Note:* Gross earnings in Euros. We adjust for inflation by using the respective CPI.

Table 5.11: Fixed effects log earnings regressions, 1997–2000

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Any training | | -.005<br>(0.009) | | | | |
| Training duration | 0.05*<br>(0.03) | 0.057*<br>(0.032) | | | | |
| Training duration full-time | | | -.188<br>(0.22) | | | |
| Training duration part-time | | | 0.361***<br>(0.087) | | | |
| Training duration less hours | | | 0.03<br>(0.037) | | | |
| Training duration correspondence course | | | -.006<br>(0.062) | | | |
| Training duration aim 1 | | | | 0.723<br>(0.816) | | |
| Training duration aim 2 | | | | -.084<br>(0.225) | | |
| Training duration aim 3 | | | | 0.023<br>(0.055) | | |
| Training duration aim 4 | | | | 0.077**<br>(0.039) | | |
| Training duration aim 5 | | | | -.006<br>(0.085) | | |
| Training duration during work | | | | | 0.166***<br>(0.057) | |
| Training duration during leisure | | | | | 0.009<br>(0.035) | |
| Duration without any financing | | | | | | 0.023<br>(0.077) |
| Duration financed by employer | | | | | | 0.06<br>(0.083) |
| Duration financed by employee | | | | | | -.002<br>(0.041) |
| Duration financed by both<br>employer and employee | | | | | | 0.183***<br>(0.064) |
| # obs. | 9,576 | 9,576 | 9,576 | 9,576 | 9,576 | 9,576 |
| $R^2$ | 0.031 | 0.031 | 0.033 | 0.031 | 0.031 | 0.031 |

*Note:* Unbalanced sample including 3,224 individuals. Training duration in years. All regressions also include a full set of year dummies. Standard errors in brackets.
Course with aim 1: retraining in a different occupation; aim 2: adjustment to standards of new job; aim 3: qualification for promotion; aim 4: adjustment to standards of current job; aim 5: other aims.
Full-time course: 30 hours per week and more; part-time courses: between 15 and 30 hours per week; less intense courses: less than 15 hours per week.

*** significant at 1%; ** significant at 5%; * significant at 10%.

Table 5.12: Fixed effects log earnings regressions: 2001–2004

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Any training | | -.012<br>(0.01) | | | | | |
| Training duration | 0.048<br>(0.036) | 0.064*<br>(0.039) | | | | | |
| Training duration full-time | | | -.047<br>(0.186) | | | | |
| Training duration part-time | | | 0.118<br>(0.127) | | | | |
| Training duration less hours | | | 0.051<br>(0.044) | | | | |
| Training duration correspondence course | | | 0.049<br>(0.105) | | | | |
| Training duration aim 1 | | | | 0.136<br>(0.147) | | | |
| Training duration aim 2 | | | | 0.04<br>(0.343) | | | |
| Training duration aim 3 | | | | 0.023<br>(0.059) | | | |
| Training duration aim 4 | | | | 0.016<br>(0.053) | | | |
| Training duration aim 5 | | | | 0.216**<br>(0.106) | | | |
| Training duration during work | | | | | -.008<br>(0.064) | | |
| Training duration during leisure | | | | | 0.076*<br>(0.044) | | |
| Duration without any financing | | | | | | -.009<br>(0.133) | |
| Duration financed by employer | | | | | | 0.0005<br>(0.086) | |
| Duration financed by employee | | | | | | 0.054<br>(0.053) | |
| Duration financed by both employer and employee | | | | | | 0.062<br>(0.075) | |
| Duration general training | | | | | | | -.037<br>(0.121) |
| Duration specific training | | | | | | | 0.056<br>(0.038) |
| # obs. | 13,728 | 13,728 | 13,728 | 13,728 | 13,728 | 13,728 | 13,728 |
| $R^2$ | 0.018 | 0.018 | 0.018 | 0.019 | 0.018 | 0.018 | 0.018 |

*Note:* Unbalanced sample including 4,487 individuals. Training duration in years. All regressions also include a full set of year dummies. Standard errors in brackets.
Course with aim 1: retraining in a different occupation; aim 2: adjustment to standards of new job; aim 3: qualification for promotion; aim 4: adjustment to standards of current job; aim 5: other aims.
Full-time course: 30 hours per week and more; part-time courses: between 15 and 30 hours per week; less intense courses: less than 15 hours per week.

*** significant at 1%; ** significant at 5%; * significant at 10%.

Table 5.13: Fixed growth rates log earnings regressions: 1997–2000

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Any training | | -.075*** | | | | |
| | | (0.025) | | | | |
| Training duration | 0.049 | 0.138 | | | | |
| | (0.09) | (0.094) | | | | |
| Training duration full-time | | | -.106 | | | |
| | | | (0.455) | | | |
| Training duration part-time | | | -.116 | | | |
| | | | (0.319) | | | |
| Training duration less hours | | | -.019 | | | |
| | | | (0.114) | | | |
| Training duration correspondence course | | | 0.302* | | | |
| | | | (0.18) | | | |
| Training duration aim 1 | | | | -.464 | | |
| | | | | (1.373) | | |
| Training duration aim 2 | | | | 0.322 | | |
| | | | | (0.568) | | |
| Training duration aim 3 | | | | 0.141 | | |
| | | | | (0.18) | | |
| Training duration aim 4 | | | | 0.004 | | |
| | | | | (0.113) | | |
| Training duration aim 5 | | | | 0.065 | | |
| | | | | (0.299) | | |
| Training duration during work | | | | | 0.166 | |
| | | | | | (0.17) | |
| Training duration during leisure | | | | | 0.006 | |
| | | | | | (0.106) | |
| Duration without any financing | | | | | | 0.536** |
| | | | | | | (0.223) |
| Duration financed by employer | | | | | | -.028 |
| | | | | | | (0.233) |
| Duration financed by employee | | | | | | -.039 |
| | | | | | | (0.129) |
| Duration financed by both employer and employee | | | | | | -.063 |
| | | | | | | (0.195) |
| # obs. | 7,182 | 7,182 | 7,182 | 7,182 | 7,182 | 7,182 |
| $R^2$ | 0.009 | 0.011 | 0.009 | 0.009 | 0.009 | 0.009 |

*Note:* Unbalanced sample including 3,224 individuals. The models are estimated by applying fixed effects to the differenced equation. Training duration in years. All regressions also include a full set of year dummies. Standard errors in brackets.
Course with aim 1: retraining in a different occupation; aim 2: adjustment to standards of new job; aim 3: qualification for promotion; aim 4: adjustment to standards of current job; aim 5: other aims.
Full-time course: 30 hours per week and more; part-time courses: between 15 and 30 hours per week; less intense courses: less than 15 hours per week.

*** significant at 1%; ** significant at 5%; * significant at 10%.

Table 5.14: Fixed growth rates log earnings regressions: 2001–2004

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Any training | | 0.016<br>(0.027) | | | | | |
| Training duration | -.018<br>(0.112) | -.036<br>(0.117) | | | | | |
| Training duration full-time | | | -.114<br>(0.432) | | | | |
| Training duration part-time | | | -.304<br>(0.426) | | | | |
| Training duration less hours | | | -.040<br>(0.135) | | | | |
| Training duration correspondence course | | | 0.322<br>(0.31) | | | | |
| Training duration aim 1 | | | | 0.359<br>(0.518) | | | |
| Training duration aim 2 | | | | 0.008<br>(0.854) | | | |
| Training duration aim 3 | | | | -.094<br>(0.196) | | | |
| Training duration aim 4 | | | | 0.06<br>(0.157) | | | |
| Training duration aim 5 | | | | -.259<br>(0.343) | | | |
| Training duration during work | | | | | -.065<br>(0.169) | | |
| Training duration during leisure | | | | | 0.019<br>(0.147) | | |
| Duration without any financing | | | | | | -.021<br>(0.295) | |
| Duration financed by employer | | | | | | 0.048<br>(0.219) | |
| Duration financed by employee | | | | | | -.062<br>(0.191) | |
| Duration financed by both employer and employee | | | | | | 0.041<br>(0.229) | |
| Duration general training | | | | | | | 0.091<br>(0.324) |
| Duration specific training | | | | | | | -.031<br>(0.119) |
| # obs. | 10,296 | 10,296 | 10,296 | 10,296 | 10,296 | 10,296 | 10,296 |
| $R^2$ | 0.004 | 0.004 | 0.004 | 0.004 | 0.004 | 0.004 | 0.004 |

*Note:* Unbalanced sample including 4,487 individuals. The models are estimated by applying fixed effects to the differenced equation. Training duration in years. All regressions also include a full set of year dummies. Standard errors in brackets.
Course with aim 1: retraining in a different occupation; aim 2: adjustment to standards of new job; aim 3: qualification for promotion; aim 4: adjustment to standards of current job; aim 5: other aims.
Full-time course: 30 hours per week and more; part-time courses: between 15 and 30 hours per week; less intense courses: less than 15 hours per week.

*** significant at 1%; ** significant at 5%; * significant at 10%.

# Chapter 6

# Concluding Remarks

*This chapter summarizes the main findings of this book. Policy implications which can be derived from these results are highlighted. Finally, an outlook for future research is provided, in which also potential shortcomings and problems are discussed.*

# Main Findings and Policy Implications

Training and lifelong learning are considered to be one solution to the problems of high unemployment rates among low-skilled workers, skill shortages, skill-biased technological change, and a lower relative demand for low-skilled labor due to international outsourcing. This book is a contribution to the ongoing debate about these issues. It studies the effects of training and lifelong learning both in a developed country (Germany) and in a transition economy (Serbia). In the latter case, it contributes to the relatively scarce literature on the effect of human capital investments in countries passing through a transitional period. Unemployment rates are high and also very persistent in Serbia; and also after the democratic changes in 2000, unemployment has further increased. On the other hand, this book focuses on Germany, a developed and industrialized economy, where nonetheless the risk of unemployment is remarkably high among low-skilled and unskilled individuals and the gap between skill-specific unemployment rates is relatively large by international standards. The empirical studies which are part of this book are based on administrative data from the Federal Employment Agency (FEA) in Germany, a special survey conducted in Serbia, and the German Socio-Economic Panel Study (SOEP). The empirical approaches mainly rely on matching estimators, but also panel estimators are applied.

Do prime-age skilled unemployed benefit more from training? Chapter 2 addresses this question as the treatment effects of public training programs for the unemployed in Germany are studied. More specifically, the picture that has been sketched in previous studies is extended by estimating treatment effects for different sub-groups of the unemployed with respect to vocational education and age. Only little evidence is found supporting heterogeneous treatment effects; and the magnitude of the differences which are found is quite small. These results are thus conflicting with the strategy to increasingly provide training to individuals with better employment prospects. This strategy has been implemented in Germany as a part of the reform of active labor market policy in 2003. After the reform, caseworkers are asked to evaluate the employment prospects of the unemployed in advance and to provide training only to individuals with a relatively high probability of entering employment after training participation. While this approach does not take into account the relative gain compared to the situation without training, the findings presented here underline the importance of doing so.

The Hartz reform in Germany introduced training vouchers and imposed more selective criteria on applicants and programs. Although it has been previously shown that the overall impact of the reform on the effectiveness of public training programs was positive, the question remains which features of the reform caused this increase—and to what particular extent. Chapter 3 isolates the effect induced by changes in the composition of program participants from the effect based on the introduction of vouchers. We find a slightly positive impact of the reform. The decomposition of this overall effect shows that the selection effect is—if at all—slightly negative. Furthermore, we find evidence that the voucher effect increased both the employment probability and earnings of the participants. This effect becomes substantially positive after around 6 months of training, and decreases slightly at the end of our observation period (1.5 years after program entry). Our results are mainly driven by skilled participants. We do not find any significant reform effect for the unskilled. While the former group can take advantage from an increased consumer sovereignty, unskilled individuals seem to have problems in adequately using the innovative voucher scheme.

Chapter 4 contributes to the still relatively scarce literature analyzing the effectiveness of active labor market policy in transition economies. More specifically, the causal impacts of participation in the *Beautiful Serbia* program are studied which provides training and temporary work in the construction sector. This evaluation deviates from routine program evaluation by considering subjective measures of individual well-being as possible outcomes. Hence, this chapter is linked to the rising economic literature focusing on the concept of happiness as an approximation for the individual welfare scale. Given that the ultimate goal of social policies is the improvement of individual welfare, subjective well-being clearly is a relevant dimension for a full impact assessment of an active labor market program. The results indeed provide an example that the positive effects of a policy can appear stronger if it is judged by subjective well-being rather than by labor market effects. The program probably impacted on individual welfare through other channels than the immediate economic status, notably by strengthening self-confidence, job desire and social inclusion of the participants. It thus appears to be important to take into account the demands and requirements of a rigorous evaluation exercise from a very early stage of the program's implementation process. This includes the appropriate design and systematic documentation of the allocation procedure and data collection.

The preceding analyses are complemented in Chapter 5 as training activities of employed individuals in Germany are studied. While the findings with respect to the wage effects of private-sector training are relatively unstable, there appear to be clearly positive effects on subsequent employment. These effects disappear after around 5 years, and seem to be solely based on whether or not an individual engaged in training at all. The respective duration which has been spent in training does not appear to matter in this context. These findings are consistent with the signaling theory (Spence, 1973), but hardly compatible with the theory of human capital (Becker, 1962, 1993).

## Future Research

This book is part of a large and still growing literature of empirical research on the evaluation of active labor market policies and private-sector training. In a broader sense, this book deals with the issue of human capital accumulation over the life cycle, or lifelong learning.

The findings of this book indicate that the introduction of vouchers in the course of the Hartz reform increased the effectiveness of public training programs. This impact is identified although the administrative data used in this book does neither include who actually received a voucher nor when he or she did so. This information, however, would open numerous avenues for future research. Besides being able to analyze the selection process into the programs more thoroughly, it would be possible to assess the effects of the introduction of the voucher in a broader dimension, such as for example its impacts in the context of a more general framework of job search.

Furthermore, the results presented in this book contribute to the relatively scarce literature analyzing the effectiveness of active labor market policy in transition economies. Lehmann and Kluve (2008) provide an overview about some studies which have been undertaken in Central and Eastern European countries in recent years. The authors also point out lessons which can be drawn from these studies. One of these lessons is that transplanting programs developed in mature market economies to transition countries has to be rather carefully done. Moreover, the evaluation approach should thoroughly take into account the specific characteristic of these countries. Importantly, the rapidly changing nature of the economic environment has to be controlled for. The authors conclude their survey by highlighting that data collection and evaluation need to be intensified before one can finally judge the efficacy of active labor market measures in transition economies—a view which is strongly supported in the light of the study included in this book.

This book focuses on the effectiveness of active labor market policy. Whereas this dimension is extensively studied in the literature, there are relatively few examples of evaluation studies incorporating a rigorous cost-benefit analysis. Notable exceptions include *inter alia* Bell and Orr (1994), Bloom et al. (1997), and Jespersen et al. (2008). However, conducting cost-benefit analysis is considered to be crucial for evidence-based policy making, which is based on facts rather than on theory or ideology (Kluve and Schmidt, 2002). For instance, in Germany where active labor market policy was systematically evaluated after the Hartz reforms, it is evident that only 28 percent of expenditures cause positive effects (Eichhorst and Zimmermann, 2007). So this share can be regarded as an upper bound in terms of efficiency—but clear evidence is still scarce. Moreover, macro-econometric studies point at considerable deadweight losses and substitution effects.

*Why* are *which* programs *when* effective for *whom*? Although this book addresses at least some aspects of this very fundamental and global question, there remain a number of avenues for future research in this direction. Obviously one needs the appropriate for doing so, which is usually not commonly available. However, the *IZA Evaluation Data Set* is supposed to provide numerous ways to study related questions (see Caliendo et al., 2009, for details). The data collection process is still ongoing in which an inflow sample of unemployed in Germany will be followed over time. The data comprises both administrative data and survey information. While the administrative data is very similar to the data used in Chapters 2 and 3 of this book, the survey information is rather novel, in particular in the context of program evaluation. Respondents answered an extensive set of questions *inter alia* about their search behavior, reservation wages, previous employment experience, and expectations about program participation. Although the currently available data do not allow for a confrontation with actual outcomes of program participation, unemployment duration, or characteristics of the accepted job, van den Berg et al. (2009) provide an example for the large potential of the *IZA Evaluation Data Set* and the perspectives it opens up—already at this relatively early stage.

Training activities of employed individuals are studied in Chapter 5 of this book. In this context, non-competitive theories of training predict that—under certain conditions—productivity growth should exceed wage growth after training (Acemoglu and Pischke, 1999). Hence, the measurement of the overall effect of training would require direct measures of productivity. Linked employer-employee data sets provide great opportunities for further research in this direction. Moreover, future research may be able to shed more light on the optimal mix between schooling, initial training, subsequent training and lifelong learning. This question has not been answered satisfactorily in the literature so far.

In theory, investments decisions in human capital are assumed to maximize the agent's lifetime income. Dynamic structural models are able to capture such considerations, and thus reveal useful insights. Additionally, this type of models are suited to analyze also the longer-term effects of interventions. Adda et al. (2006) is a good example for such an approach to evaluate the life cycle return to apprenticeship training. The type of model is similar to the models used to analyze schooling decisions which are rather intensively studied in the literature (Keane and Wolpin, 1997; Eckstein and Wolpin, 1999; Belzil and Hansen, 2002).

Finally, the debate which has been briefly mentioned in the introduction to this book still needs to be answered satisfactorily by future research. It is precisely the question whether or not (and if so, under which circumstances) Becker's approach towards education as increasing a worker's productivity dominates Spence's view on education as also—and perhaps primarily—serving as a signal for otherwise unobservable abilities. To be able to answer this question, it is clearly important to take into account the endogenous nature of training or education decisions very thoroughly.

# Bibliography


Abadie, A. and G. W. Imbens (2002). Simple and Bias-Corrected Matching Estimators for Average Treatment Effects. NBER Technical Working Paper 286, National Bureau of Economic Research, Inc, Cambridge, MA.

Abadie, A. and G. W. Imbens (2008). On the Failure of the Bootstrap for Matching Estimators. *Econometrica 76*(6), 1537–1557.

Acemoglu, D. (2002). Technical Change, Inequality, and the Labor Market. *Journal of Economic Literature 40*(1), 7–72.

Acemoglu, D. and J.-S. Pischke (1999). Beyond Becker: Training in Imperfect Labour Markets. *The Economic Journal 109*(453), F112–F142.

Adda, J., C. Dustmann, C. Meghir, and J.-M. Robin (2006). Career Progression and Formal versus On-the-Job Training. IZA Discussion Paper 2260, Institute for the Study of Labor (IZA), Bonn.

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.

Arandarenko, M. (2004). International Advice and Labor Market Institutions in South-East Europe. *Global Social Policy 4*(1), 27–53.

Arandarenko, M. and M. Jovicic (2007). Regional Labour Market Differences in Serbia: Assessment and Policy Recommendations. *The European Journal of Comparative Economics 4*(2), 299–317.

Barnow, B. S. (2000). Vouchers for Federal Targeted Training Programs. In C. E. Steuerle, V. Doorn Ooms, G. E. Peterson, and R. D. Reischauer (Eds.), *Vouchers and the Provision of Public Services*. Washington D.C.: Brookings Institution Press.

Barnow, B. S. (2009). Vouchers in U.S. Vocational Training Programs: An Overview of What We Have Learned. *Journal for Labour Market Research 42*(1), 71–84.

Bauer, T. K. and S. Bender (2004). Technological Change, Organizational Change, and Job Turnover. *Labour Economics 11*(3), 265–291.

Büchel, F. and M. Pannenberg (2004). Berufliche Weiterbildung in West- und Ostdeutschland: Teilnehmer, Struktur und individueller Ertrag. *Journal for Labour Market Research 37*(2), 73–126.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy 70*(5), 9–49.

Becker, G. S. (1993). *Human Capital: A Theoretical and Empirical Analysis with a Special Reference to Education* (3rd ed.). Chicago: University of Chicago Press.

Bell, S. H. and L. L. Orr (1994). Is subsidized employment cost effective for welfare recipients? experimental evidence from seven state demonstrations. *Journal of Human Resources 29*(1), 42–61.

Belzil, C. and J. Hansen (2002). Unobserved Ability and the Return to Schooling. *Econometrica 70*(5), 2075–2091.

Berman, E., J. Bound, and S. Machin (1998). Implications of Skill-Biased Technological Change: International Evidence. *Quarterly Journal of Economics 113*(4), 1245–1279.

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, Institute for the Study of Labor (IZA), Bonn.

Black, J. (2002). *A Dictionary of Economics* (2nd ed.). Oxford and New York: Oxford University Press.

Bloom, H. S., L. L. Orr, S. H. Bell, G. Cave, F. Doolittle, W. Lin, and J. M. Bos (1997). The benefits and costs of jtpa title ii-a programs: Key findings from the national job training partnership act study. *Journal of Human Resources 32*(3), 549–576.

Bonin, H. and U. Rinne (2006). Beautiful Serbia. IZA Discussion Paper 2533, Institute for the Study of Labor (IZA), Bonn.

Bonin, H. and H. Schneider (2006). Workfare: Eine wirksame Alternative zum Kombilohn. *Wirtschaftsdienst 86*(10), 645–650.

Bonin, H., M. Schneider, H. Quinke, and T. Arens (2007). Zukunft von Bildung und Arbeit. IZA Research Report 9, Institute for the Study of Labor (IZA), Bonn.

Brenke, K. and K. F. Zimmermann (2005). Demographic Change Necessitates Educational Reform and Lifelong Learning. DIW Weekly Report 17/2005, German Institute for Economic Research (DIW), Berlin.

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., A. Falk, L. C. Kaiser, H. Schneider, A. Uhlendorff, G. J. van den Berg, and K. F. Zimmermann (2009). The IZA Evaluation Data Set – The First Wave. Working Paper, Institute for the Study of Labor (IZA), Bonn.

Caliendo, M., R. Hujer, and S. L. Thomsen (2008). Identifying Effect Heterogeneity to Improve the Efficiency of Job Creation Schemes in Germany. *Applied Economics 40*(9), 1101–1122.

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. *Journal for Labour Market Research 38*(2-3), 396–418.

Christensen, B. (2001). Berufliche Weiterbildung und Arbeitsplatzrisiko: Ein Matching-Ansatz. Kiel Working Paper 1033, Kiel Institute for the World Economy (IfW), Kiel.

Clark, A. E., P. Frijters, and M. A. Shields (2006). Income and Happiness: Evidence, Explanations and Economic Implications. PSE Discussion Paper 2006-24, PSE (Ecole normale supérieure).

Clark, A. E., P. Frijters, and M. A. Shields (2008). Relative Income, Happiness and Utility: An Explanation for the Easterlin Paradox and Other Puzzles. *Journal of Economic Literature 46*(1), 95–144.

Dearden, L., H. Reed, and J. van Reenen (2000). Who Gains when Workers Train? Training and Corporate Productivity in a Panel of British Industries. CEPR Discussion Paper 2486, Centre for Economic Policy Research (CEPR), London.

Di Tella, R. and R. J. MacCulloch (2006). Some Uses of Happiness Data in Economics. *The Journal of Economic Perspectives 20*(1), 25–46.

Di Tella, R., R. J. MacCulloch, and A. J. Oswald (2003). The Macroeconomics of Happiness. *The Review of Economics and Statistics 85*(4), 809–827.

Dustmann, C. (2004). Parental Background, Secondary School Track Choice, and Wages. *Oxford Economic Papers 56*(2), 209–230.

Dustmann, C. and C. Meghir (2005). Wages, Experience and Seniority. *The Review of Economic Studies 72*(1), 77–108.

Ebbinghaus, B. and W. Eichhorst (2009). Employment Regulation and Labor Market Policy in Germany, 1991–2005. In P. de Beer and T. Schils (Eds.), *Social Policy and the Labour Market: Achieving an Optimal Policy Mix*. Cheltenham: Edward Elgar.

Eckstein, Z. and K. I. Wolpin (1999). Why Youths Drop Out of High School: The Impact of Preferences, Opportunities, and Abilities. *Econometrica 67*(6), 1295–1340.

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.

Evertsson, M. (2004). Formal on-the-job training: A gender-typed experience and wage-related advantage? *European Sociologocal Review 20*(1), 79–94.

Feenstra, R. C. and G. Hanson (2001). Global Production Sharing and Rising Inequality: A Survey of Trade and Wages. NBER Working Paper 8372, National Bureau of Economic Research, Inc, Cambridge, MA.

Fitzenberger, B., A. Osikominu, and R. Völter (2008). Get Training or Wait? Long-Run Employment Effects of Training Programs for the Unemployed in West Germany. *Annales d'Economie et de Statistique* (forthcoming).

Frazis, H. and M. A. Loewenstein (2005). Reexamining the Returns to Training: Functional Form, Magnitude, and Interpretation. *Journal of Human Resources 40*(2), 453–476.

Frey, B. and A. Stutzer (2000). Happiness, Economy, and Institutions. *The Economic Journal 110*(466), 918–938.

Frey, B. and A. Stutzer (2002). *Happiness and Economics: How the Economy and Institutions Affect Human Well-Being*. Princeton and Oxford: Princeton University Press.

Friedman, M. (1962). The Role of Government in Education. In *Capitalism and Freedom*, Chapter 6. Chicago: University of Chicago Press.

Geishecker, I. (2008). The Impact of International Outsourcing on Individual Employment Security: A Micro-Level Analysis. *Labour Economics 15*(3), 291–314.

Geishecker, I. and H. Görg (2008). Winners and Losers: A Micro-Level Analysis of International Outsourcing and Wages. *Canadian Journal of Economics 41*(1), 243–270.

Gruber, J. and S. Mullainathan (2005). Do Cigarette Taxes Make Smokers Happier? *Advances in Economic Analysis & Policy 5*(1), Article 4.

Hayo, B. (2004). Happiness in Eastern Europe. Marburg Working Papers on Economics 200412, Philipps-Universität Marburg, Faculty of Business Administration and Economics, Department of Economics.

Hayo, B. and W. Seifert (2003). Subjective Economic Well-Being in Eastern Europe. *Journal of Economic Psychology 24*(3), 329–348.

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., C. Heinrich, and J. A. Smith (1997). Assessing the Performance of Performance Standards in Public Bureaucracies. *American Economic Review 87*(2), 389–395.

Heckman, J. J. and V. J. Hotz (1989). Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training. *Journal of the American Statistical Association 84*(408), 862–874.

Heckman, J. J., H. Ichimura, and P. E. Todd (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies 64*(4), 605–654.

Heckman, J. J., H. Ichimura, and P. E. Todd (1998). Matching as an Econometric Evaluation Estimator. *The Review of Economic Studies 65*(2), 261–294.

Heckman, J. J., R. J. LaLonde, and J. A. Smith (1999). The Economics and Econometrics of Active Labor Market Programs. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3, Chapter 31, pp. 1865–2097. Amsterdam: Elsevier.

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 Arbeitslosigkeitsdauer in West- und Ostdeutschland: Eine mikroökonometrische Analyse. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung 33*(3), 405–419.

Hummels, D. L., J. Ishii, and K.-M. Yi (2001). The Nature and Growth of Vertical Specialization in World Trade. *Journal of International Economics 54*(1), 75–96.

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. and J. M. Wooldridge (2008). Recent Developments in the Econometrics of Program Evaluation. IZA Discussion Paper 3640, Institute for the Study of Labor (IZA), Bonn.

Jacobebbinghaus, P. and S. Seth (2007). The German Integrated Employment Biographies Sample IEBS. *Schmollers Jahrbuch 127*(2), 335–342.

Jacobi, L. and J. Kluve (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.

Jespersen, S. T., J. R. Munch, and L. Skipper (2008). Costs and Benefits of Danish Active Labour Market Programmes. *Labour Economics 15*(5), 859–884.

Kailis, E. and S. Pilos (2005). Lifelong Learning in Europe. *Eurostat: Statistics in Focus 8/2005.*

Keane, M. P. and K. Wolpin (1997). The Career Decisions of Young Men. *Journal of Political Economy 105*(3), 473–522.

Kluve, J. (2006). The Effectiveness of European Active Labor Market Policy. IZA Discussion Paper 2018, Institute for the Study of Labor (IZA), Bonn.

Kluve, J., H. Lehmann, and C. M. Schmidt (1999). Active Labor Market Policies in Poland: Human Capital Enhancement, Stigmatisation, or Benefit Churning? *Journal of Comparative Economics 27*, 61–89.

Kluve, J. and C. M. Schmidt (2002). Can Training and Employment Subsidies Combat European Employment? *Economic Policy 35*, 411–448.

Korpi, T. (1997). Is Utility Related to Employment Status? Employment, Unemployment, Labor Market Policies and Subjective Well-Being Among Swedish Youth. *Labour Economics 4*(2), 125–147.

Krueger, A. B. and P. Zhu (2004). Another Look at the New York City School Voucher Experiment. *American Behavioral Scientist 47*(5), 658–698.

Kruppe, T. (2008). Selektivität bei der Einlösung von Bildungsgutscheinen. IAB Discussion Paper 17/2008, Institute for Employment Research (IAB), Nuremberg.

Kuwan, H. and F. Thebis (2004). Berichtssystem Weiterbildung IX: Ergebnisse der Repräsentativbefragung zur Weiterbildungssituation in Deutschland. Federal Ministry of Education and Research (BMBF), Berlin.

Ladd, H. F. (2002). School Vouchers: A Critical Review. *Journal of Economic Perspectives 16*(4), 3–24.

LaLonde, R. J. (2003). Employment and Training Programs. In M. Feldstien and R. Moffitt (Eds.), *Means Tested Transfer Programs in the U.S.* Chicago: University of Chicago Press for the National Bureau of Economic Research.

Layard, R. (2006). Happiness and Public Policy: A Challenge to the Profession. *The Economic Journal 116*(510), C24–33.

Lechner, M. (1999). Earnings and Employment Effects of Continuous Off-the-Job Training in East-Germany after Unification. *Journal of Business and Economic Statistics 17*(1), 74–90.

Lechner, M. (2000). An Evaluation of Public-Sector-Sponsored Continuous Vocational Training Programs in East Germany. *Journal of Human Resources 35*(2), 347–375.

Lechner, M. (2001). Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption. In M. Lechner and F. Pfeiffer (Eds.), *Econometric Evaluation of Labour Market Policies*, pp. 1–18. Heidelberg: Physica-Verlag.

Lechner, M. (2002). Some Practical Issues in the Evaluation of Heterogeneous Labour Market Programmes by Matching Methods. *Journal of the Royal Statistical Society, Series A, Statistics in Society 165*(1), 59–82.

Lechner, M. and B. Melly (2007). Earnings Effects of Training Programs. IZA Discussion Paper 2926, Institute for the Study of Labor (IZA), Bonn.

Lechner, M., R. Miquel, and C. Wunsch (2005). Long-Run Effects of Public Sector Sponsored Training in West Germany. IAB Discussion Paper 3/2005, Institute for Employment Research (IAB), Nuremberg.

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 (2006a). Active Labour Market Policy in East Germany: Waiting for the Economy to Take Off. IZA Discussion Paper 2363, Institute for the Study of Labor (IZA), Bonn.

Lechner, M. and C. Wunsch (2006b). Are Training Programs More Effective When Unemployment Is High? IZA Discussion Paper 2355, Institute for the Study of Labor (IZA), Bonn.

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.

Leetmaa, R. and A. Võrk (2004). Evaluation of Active Labour Market Programmes in Estonia. *mimeo*.

Lehmann, H. and J. Kluve (2008). Assessing Active Labor Market Policies in Transition Economies. Working Paper 646, University of Bologna, Department of Economics, Bologna.

Leuven, E. (2005). The Economics of Private-Sector Training: A Survey of the Literature. *Journal of Economic Surveys 19*(1), 91–111.

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.

Lubyova, M. and J. C. van Ours (1999). Effects of Active Labor Market Programs on the Transition Rate from Unemployment into Regular Jobs in the Slovak Republic. *Journal of Comparative Economics 27*(1), 90–112.

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.

McQuaid, R. W. and C. Lindsay (2005). The Concept of Employability. *Urban Studies 42*(2), 197–219.

Melis, A. (2006). Expenditure on Training Measures for the Unemployed across the EU. *Eurostat: Statistics in Focus 5/2006*.

Micklewright, J. and G. Nagy (2005). Job Search Monitoring and Unemployment Duration in Hungary: Evidence from a Randomised Control Trial. IZA Discussion Paper 1839, Institute for the Study of Labor (IZA), Bonn.

Mincer, J. (1962). On-the-Job Training: Costs, Returns and Some Implications. *Journal of Political Economy 70*(5), S50–S79.

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.

Neyman, J. S. (1923). On the Application of Probability Theory to Agriculture Experiments. Essay on Principles. Section 9. *Roczniki Nauk Rolniczych Tom X*, 1–51 (Annals of Agricultural Sciences). Translated in: J. S. Neyman, D. M. Dabrowska and T. P. Speed (1990), *Statistical Science 5*(4), 465–472.

OECD (2006). *Employment Outlook. Boosting Jobs and Incomes.* Paris: Organization for Economic Cooperation and Development.

O'Leary, C. J. (1998a). Evaluating the Effectiveness of Active Labor Market Programs in Hungary. Upjohn Institute Technical Report 98-013, W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.

O'Leary, C. J. (1998b). Evaluating the Effectiveness of Active Labor Market Programs in Poland. Upjohn Institute Technical Report 98-012, W.E. Upjohn Institute for Employment Research, Kalamazoo, Michigan.

Pfeiffer, F. (2001). Training and Individual Performance in Europe: Evidence from Microeconometric Studies. In P. Descy and M. Tessaring (Eds.), *Training in Europe*, pp. 7–41. Luxembourg.

Pigou, A. C. (1912). *Wealth and Welfare.* London: Macmillan.

Pischke, J.-S. (2001). Continuous Training in Germany. *Journal of Population Economics 14*(3), 523–548.

Puhani, P. (2002). Advantage through Training in Poland? A Microeconometric Evaluation of the Employment Effects of Training and Job Subsidy Programmes. *Labour 16*(3), 569–608.

Puhani, P. and V. Steiner (1997). The Effectiveness and Efficiency of Active Labour Market Policies in Poland. *Empirica 24*(3), 209–231.

Reinberg, A. and M. Hummel (2005). Höhere Bildung schützt auch in der Krise vor Arbeitslosigkeit. IAB Kurzbericht 9/2005, Institute for Employment Research (IAB), Nuremberg.

Reinberg, A. and M. Hummel (2007). Der Trend bleibt – Geringqualifizierte sind häufiger arbeitslos. IAB Kurzbericht 18/2007, Institute for Employment Research (IAB), Nuremberg.

Rinne, U., M. Schneider, and A. Uhlendorff (2007). Too Bad to Benefit? Effect Heterogeneity of Public Training Programs. IZA Discussion Paper 3240, Institute for the Study of Labor (IZA), Bonn.

Rinne, U., A. Uhlendorff, and Z. Zhao (2008). Vouchers and Caseworkers in Public Training Programs: Evidence from the Hartz Reform in Germany. IZA Discussion Paper 3910, Institute for the Study of Labor (IZA), Bonn.

Rodriguez-Planas, N. (2007). What Works Best for Getting the Unemployed Back to Work: Employment Services or Small-Business Assistance Programmes? Evidence from Romania. IZA Discussion Paper 3051, Institute for the Study of Labor (IZA), Bonn.

Rodriguez-Planas, N. and J. Benus (2009). Evaluating Active Labor Market Programs in Romania. *Empirical Economics* (forthcoming).

Rosen, S. (1998). Human Capital. In J. Eatwell, M. Milgate, and P. K. Newman (Eds.), *The New Palgrave: A Dictionary of Economics*, Volume 2, pp. 681–690. London: Macmillan.

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.

Rosenbaum, P. R. and D. B. Rubin (1985). Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. *The American Statistican 39*(1), 33–38.

Rosenstein-Rodan, P. (1943). Problem of Industrialization of Eastern and Southern-Eastern Europe. *The Economic Journal 53*, 202–211.

Roy, A. D. (1951). Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers 3*, 135–146.

Rubin, D. B. (1973). The Use of Matched Sampling and Regression Adjustment to Remove Bias in Observational Studies. *Biometrics 29*, 185–203.

Rubin, D. B. (1974). Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies. *Journal of Educational Psychology 66*(5), 688–701.

Rutherford, D. (1992). *Dictionary of Economics*. London and New York: Routledge.

Sauermann, J. (2006). Who Invests in Training if Contracts are Temporary? Empirical Evidence for Germany Using Selection Correction. IWH Discussion Paper 14/2006, Halle Institute for Economic Research (IWH), Halle (Saale).

Schneider, F. (2007). Shadow Economies and Corruption All Over the World: New Estimates for 145 Countries. *Economics – The Open-Access, Open-Assessment E-Journal 1*(9).

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, Institute for the Study of Labor (IZA), Bonn.

Schneider, H. and A. Uhlendorff (2005). The Transition from Welfare to Employment and the Role of Potential Labor Income. *Journal of Applied Social Science Studies 125*(1), 51–61.

Schneider, H. and A. Uhlendorff (2006). Die Wirkung der Hartz-Reform im Bereich der beruflichen Weiterbildung. *Journal for Labour Market Research 39*(3-4), 477–490.

Schnur, P. and G. Zika (2007). Die Grenzen der Expansion. IAB Kurzbericht 26/2007, Institute for Employment Research (IAB), Nuremberg.

Sianesi, B. (2004). An Evaluation of the Active Labor Market Programmes in Sweden. *The Review of Economics and Statistics 86*(1), 133–155.

Smith, J. A. and P. E. Todd (2005a). Does Matching Overcome LaLonde's Critique of Non-Experimental Estimators? *Journal of Econometrics 125*(1-2), 305–353.

Smith, J. A. and P. E. Todd (2005b). Rejoinder. *Journal of Econometrics 125*(1-2), 365–375.

Spence, A. M. (1973). Job Market Signaling. *Quarterly Journal of Economics 87*(3), 355–374.

Spitz-Oener, A. (2006). Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics 24*(2), 235–270.

Steiner, V. (2004). Social Welfare Reform and the Low-Wage Labor Market in Germany – What Works, and What Doesn't? *Applied Economics Quarterly Supplement 55*, 57–78.

Stephan, G. (2008). The Effects of Active Labor Market Programs in Germany: An Investigation Using Different Definitions of Non-Treatment. *Jahrbücher für Nationalökonomie und Statistik 228*(5-6), 586–611.

Steuerle, C. E. (2000). Common Issues for Voucher Programs. In C. E. Steuerle, V. Doorn Ooms, G. E. Peterson, and R. D. Reischauer (Eds.), *Vouchers and the Provision of Public Services*. Washington D.C.: Brookings Institution Press.

Uhlendorff, A. (2008). *Transitions on the Labor Market: Unemployment, Transfer Receipt and the Low-Wage-Sector*. Saarbrücken: VDM-Verlag Dr. Müller.

van den Berg, G. J., A. Bergemann, and M. Caliendo (2009). The Effect of Active Labor Market Programs on Not-Yet Treated Unemployed Individuals. *Journal of the European Economic Association* (forthcoming).

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, Institute for the Study of Labor (IZA), Bonn.

Wagner, G. G., J. R. Frick, and J. Schupp (2007). The German Socio-Economic Panel Study (SOEP) – Scope, Evolution and Enhancements. *Schmollers Jahrbuch 127*(1), 139–169.

Wilkens, I. and U. Leber (2003). Partizipation an beruflicher Weiterbildung – Empirische Ergebnisse auf Basis des Sozio-Ökonomischen Panels. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung 36*(3), 329–337.

Winkelmann, R. (1996). Employment Prospects and Skill Acquisition of Apprenticeship-Trained Workers in Germany. *Industrial and Labor Relations Review 49*(4), 658–672.

Winterhager, H., A. Heinze, and A. Spermann (2006). Deregulating Job Placement in Europe: A Microeconometric Evaluation of an Innovative Voucher Scheme in Germany. *Labour Economics 13*(4), 505–517.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge and London: MIT Press.

Zhao, Z. (2008). Sensitivity of Propensity Score Methods to the Specifications. *Economics Letters 98*(3), 309–319.

Zimmermann, K. F. (2003). Beschäftigungspotentiale im Niedriglohnsektor. *Quarterly Journal of Economic Research 72*(1), 11–24.

# List of Tables

# List of Figures

# German Summary

## Humankapitalinvestitionen: Berufliche Weiterbildung und lebenslanges Lernen

*Die vorliegende Dissertation beschäftigt sich mit Humankapitalinvestitionen in Form von Weiterbildungsaktivitäten und lebenslangem Lernen. Dabei werden folgende Fragen beantwortet: Gibt es unter den Arbeitslosen Personengruppen, die von beruflichen Weiterbildungsmaßnahmen nicht bzw. nur eingeschränkt profitieren? Wie ist die Einführung von Bildungsgutscheinen im Zuge der Hartz-Reform zu bewerten? Wie wirkt aktive Arbeitsmarktpolitik in einem Transformationsland? Wer beteiligt sich an lebenslangem Lernen in Form von berufsbegleitenden Weiterbildungsaktivitäten und welche Effekte resultieren daraus?*

***Ausgangslage und Problemstellung***

*"Eine Investition in Wissen zahlt sich immer aus."* Diese freie Übersetzung eines Zitates von Benjamin Franklin (1706–1790) macht deutlich, dass die herausragende Bedeutung von Investitionen in Humankapital bereits früh erkannt wurde. Heutzutage wird ihnen ebenfalls eine große Bedeutung beigemessen, wobei sie zum Teil sogar als "Stein der Weisen" erscheinen, der sämtliche ökonomische Probleme und Herausforderungen der Zukunft lösen kann. So betonte etwa Barack Obama, der jetzige Präsident der Vereinigten Staaten von Amerika, in einer Ansprache am 8. Januar 2009 die Bedeutung von Bildungsinvestitionen in den Zeiten der ökonomischen Krise. Das amerikanische Konjunkturpaket umfasst Maßnahmen zur Modernisierung von Bildungseinrichtungen, und es sei damit geeignet *"die Schüler aus Chicago und Boston auf den Wettbewerb mit den Schülern aus Peking um die High-Tech Arbeitsplätze der Zukunft vorzubereiten"*. Ähnliche Strategien werden derzeit in einer Vielzahl von Ländern verfolgt.

Im Gegensatz zu Investitionen in die Schulbildung, die das Humankapital zukünftiger Generationen betreffen, haben Investitionen in die berufliche Weiterbildung eine direkte Auswirkung auf das Humankapital gegenwärtiger Erwerbspersonen. Sie stellen damit eine reizvolle Option für Politiker dar, deren Planung sich typscherweise an einem vergleichsweise kurzen Horizont orientiert. Die chinesische Regierung hat etwa kürzlich als Reaktion auf die globale Wirtschafts- und Finanzkrise bekanntgegeben, Weiterbildungmaßnahmen für Arbeitslose verstärkt zu fördern. Ähnliche Elemente finden sich auch im so genannten Konjunkturpaket II in Deutschland wieder, das sich Anfang 2009 im Gesetzgebungsprozess befindet. Insgesamt machen diese Maßnahmen deutlich, dass Humankapitalinvestitionen nach wie vor ein wichtiges Element im wirtschaftspolitischen Instrumentarium darstellen. Die vorliegende Arbeit untersucht vor diesem Hintergrund die Effektivität derartiger Investitionen.

Humankapital wird sehr allgemein als eine menschliche Produktionskapazität definiert. Investitionen in Humankapital können verschiedene Formen annehmen, die unter anderem die schulische oder berufliche Ausbildung, berufliche Weiterbildungsaktivitäten, die medizinische Versorgung oder Migration beinhalten. Diesen Aktivitäten ist gemein, dass das Wissen, die Fähigkeiten und Qualifikationen oder der Gesundheitszustand einer Person verbessert werden. Sie beeinflussen zudem in unterschiedlich starkem Ausmaß das (zukünftige) Einkommens- und Konsumprofil einer Person. Schulische oder berufliche Aus- und Weiterbildungsaktivitäten sind die bedeutendsten Formen von Humankapitalinvestitionen.

Um die Effektivität von Humankapitalinvestitionen sowohl in Industrie- als auch Transformationsländern zu untersuchen, beschäftigt sich der vorliegende Beitrag mit der Bundesrepublik Deutschland sowie Serbien. Letzteres Land befindet sich nach wie vor in einem Transformationsprozess, der mit sehr hoher Arbeitslosigkeit einhergeht. Das zentrale Erkenntnisinteresse liegt hierbei darin, inwieweit aktive Arbeitsmarktpolitik in diesem Umfeld die negative Begleiterscheinung des ökonomischen Wandels zumindest abmildern kann. Auf der anderen Seite ist das Arbeitslosigkeitsrisiko in Deutschland unter den Niedrig- und Geringqualifizierten besonders hoch. Zudem hat die Korrelation zwischen Ausbildungsniveau und Arbeitslosigkeit in der jüngsten Vergangenheit weiter zugenommen. Während in den Jahren 2004 und 2005 jeweils etwa ein Fünftel der Personen ohne Ausbildungsabschluss als arbeitslos registriert waren, lag dieser Anteil in den 1970er Jahren noch bei nur rund fünf Prozent. Der Unterschied der qualifikationsspezifischen Arbeitslosenquoten ist auch im internationalen Vergleich hoch. Vorausrechnungen von Arbeitsangebot und -nachfrage haben außerdem gezeigt, dass sich das Problem des so genannten Fachkräftemangels in Deutschland in naher Zukunft weiter zu verschärfen droht. Ein Szenario, in dem ein Mangel an qualifizierten Fachkräften mit einer hohen und persistenten Arbeitslosenquote der Niedrig- und Geringqualifizierten einhergeht, ist nach diesen Berechnungen nicht unwahrscheinlich. Verstärkte Bildungsinvestitionen könnten jedoch dieses Szenario verhindern.

Insbesondere in Deutschland wird häufig angeführt, dass die Schaffung eines Niedriglohnsektors für die Überwindung der hohen Arbeitslosigkeit unter den Niedrig- und Geringqualifizierten notwendig sei und dass das derzeitige Steuer- und Transfersystem insbesondere in diesem Bereich nur geringe Anreize zur Arbeitsaufnahme impliziere. Während diese Argumentation auf der Seite des Arbeitsangebotes ansetzt, finden andererseits Entwicklungstendenzen statt, die die Arbeitsnachfrage beeinflussen. So stellt der andauernde technische und technologische Wandel, der Individuen je nach Ausbildungsniveau unterschiedlich betrifft, eine weitere Ursache für die zunehmende Bedeutung von Humankapitalinvestitionen dar. Obgleich Befunde darauf hinweisen, dass der technologische Wandel bereits das gesamte 20. Jahrhundert Personen je nach Qualifikationsniveau in unterschiedlich starkem Ausmaß betroffen hat, gibt es auch Hinweise darauf, dass sich dieser Prozess in den letzten Jahrzehnten noch deutlich beschleunigt hat. Studien zeigen, dass dieses Phänomen auch in Deutschland anzutreffen ist. Die Zunahme von weltweiten Produktionsverlagerungen verstärken diese Tendenzen.

Lebenslanges Lernen und berufliche Weiterbildungsaktivitäten stellen eine Möglichkeit dar, den oben skizzierten Problemen auf der Arbeitsangebots- und Arbeitsnachfrageseite entgegenzuwirken. Der vorliegende Beitrag widmet sich diesem Themenkomplex und stellt somit einen Beitrag zu den anhaltenden Diskussionen dar, wie diese Herausforderungen bewältigt werden können.

### *Inhalt dieser Dissertation*

Gibt es unter den Arbeitslosen Personengruppen, die von beruflichen Weiterbildungsmaßnahmen nicht bzw. nur eingeschränkt profitieren? Kapitel 2 dieser Dissertation untersucht diese Fragestellung. Obgleich bereits eine Vielzahl früherer Untersuchungen vorliegen, die sich mit den Wirkungen der Förderung der beruflichen Weiterbildung (FbW) befassen, werden die Befunde dieser Studien erweitert, indem die potentielle Heterogenität der Maßnahmewirkungen für verschiedene Qualifikationsniveaus und Altersgruppen untersucht wird. Die Berechnungen zeigen, dass sich die Teilnahme an FbW-Maßnahmen für sämtliche untersuchte Gruppen positiv auf die künftige Beschäftigungswahrscheinlichkeit auswirkt. Ein weiterer Befund ist, dass Teilnehmer höher entlohnte Tätigkeiten auf dem ersten Arbeitsmarkt finden als vergleichbare Nichtteilnehmer. Hingegen gibt es den Berechnungen zu Folge keine Hinweise auf stärkere Heterogenitäten in den Teilnahmeeffekten. Die Unterschiede in den Effekten fallen vergleichsweise gering aus. Insbesondere stehen diese Ergebnisse damit in Widerspruch zu einer Strategie, bevorzugt Personen mit besseren Arbeitsmarktchancen eine Teilnahme an FbW-Maßnahmen zu ermöglichen.

Kapitel 3 setzt sich mit der Fragestellung auseinander, wie die Einführung von Bildungsgutscheinen im Zuge der Hartz-Reform zu bewerten ist. Frühere Studien haben bereits eine insgesamt positive Wirkung der Reform auf die Effektivität der FbW-Maßnahmen gezeigt, allerdings wurde in diesem Zusammenhang nicht der Frage nachgegangen, welche Elemente der Reform in welchem Ausmaß zu diesen Effektivitätssteigerungen geführt haben. Aus diesem Grund wird der Reformeffekt zerlegt: Einerseits in den Effekt, der auf die Veränderungen in der Zusammensetzung der Teilnehmer an FbW-Maßnahmen zurückzuführen ist, und andererseits in den Effekt, der auf die Einführung des Bildungsgutscheines zurückzuführen ist. Die Ergebnisse zeigen, dass der positive Reformeffekt, wenn überhaupt, nur in geringem Maße auf Veränderungen der Teilnehmerstruktur zurückzuführen ist. Hingegen hat die Einführung des Gutscheinsystems erhebliche Effektivitätssteigerungen hervorgerufen, die sich jedoch nicht bei den Niedrig- und Geringqualifizierten einstellen.

Wie wirkt aktive Arbeitsmarktpolitik in einem Transformationsland? Mit dieser Fragestellung setzt sich Kapitel 4 dieser Dissertation auseinander. Es werden darin die Teilnahmeeffekte an einem Programm der aktiven Arbeitsmarktpolitik analysiert, welches in Serbien implementiert wurde (*Beautiful Serbia*). Dieses Programm bestand aus zwei faktisch unabhängigen Programmteilen: Arbeitslose nahmen an Weiterbildungsmaßnahmen teil und es bestand (anschließend) die Möglichkeit, einen temporären Arbeitsplatz im Baugewerbe zu erhalten. Die Effektivität dieses Programmes wird sowohl in Hinblick auf die nachfolgende Erwerbsbiographie als auch hinsichtlich einer Reihe von Indikatoren des subjektiven Wohlbefindens ausgewertet. Interessanterweise resultiert der Befund, dass sich der Teilnahmeeffekt auf das subjektive Wohlbefinden positiver gestaltet als auf die Erwerbsbiographie.

Kapitel 5 dieses Buches geht der Fragestellung nach, wer sich an lebenslangem Lernen in Form von Weiterbildungsaktivitäten im Beruf beteiligt und welche Effekte daraus auf die Lohnentwicklung sowie auf die zukünftige Erwerbsbiographie resultieren. Auf Grundlage von Daten für zwei verschiedene Perioden (von 1997 bis 2000 sowie von 2001 bis 2004) zeigen die Untersuchungen der Beteiligung an Weiterbildungsaktivitäten ein sehr ähnliches Bild für beide Untersuchungszeiträume. Auf der anderen Seite ergibt sich ein recht unklares Bild hinsichtlich der Effekte auf die individuelle Lohnentwicklung, während sich eine Teilnahme an berufsbegleitender Weiterbildung positiv auf die Wahrscheinlichkeit einer zukünftigen Erwerbstätigkeit auswirkt. Dieser Effekt ist bis etwa fünf Jahre nach dem Ende der Weiterbildungsaktivitäten zu beobachten.

### *Schlussfolgerungen und politische Handlungsempfehlungen*

Auf Grundlage der Ergebnisse, die in dieser Dissertation dargestellt werden, lassen sich eine Reihe von Schlussfolgerungen und politische Handlungsempfehlungen ableiten. So zeigen die Resultate aus Kapitel 2 hinsichtlich der Effektheterogenität der Teilnahmeeffekte an FbW-Maßnahmen, dass die Effekte über verschiedene Qualifikationsniveaus und Altersgruppen relativ ähnlich sind. Damit ergibt sich ein Widerspruch zu der Strategie, bevorzugt Personen mit besseren Arbeitsmarktchancen eine Teilnahme an FbW-Maßnahmen zu ermöglichen, welche nach Inkrafttreten eines Paketes der Hartz-Reformen ab dem 1. Januar 2003 verfolgt wird. Den Ergebnissen dieser Arbeit zu Folge ist durch dieses Vorgehen keine Verbesserung der Effektivität der FbW-Maßnahmen zu erwarten.

Kapitel 3 beschäftigt sich eingehender mit der soeben beschriebenen Reform. Die These, dass Veränderungen in der Teilnehmerstruktur die Effektivität der FbW-Maßnahmen nur geringfügig beeinflusst hat, wird in diesem Zusammenhang empirisch bestätigt. Darüber hinaus zeigen weitere Ergebnisse, dass die Einführung des Bildungsgutscheines als innovatives Allokationsinstrument zu zum Teil erheblichen Steigerungen der Effektivität geführt hat. Es finden sich jedoch auch Hinweise darauf, dass die Gruppe der Niedrig- und Geringqualifizierten Probleme mit der Handhabung dieses Instrumentes hat und nicht von der Reform profitiert.

Die Evaluation eines speziellen Programmes der aktiven Arbeitsmarktpolitik in Serbien in Kapitel 4 weicht insofern vom Standard der Vorgehensweise ab, als dass neben der Wirkung auf die nachfolgende Erwerbsbiographie auch der Effekt auf eine Reihe von Indikatoren des subjektiven Wohlbefindens analysiert wird. Die Ergebnisse zeigen, dass die Effekte einer Maßnahme auf Grundlage des subjektiven Wohlbefindens positiver ausfallen können als auf Grundlage der Erwerbsbiographie. Dies deutet darauf hin, dass das untersuchte Programm das subjektive Wohlbefinden über andere Kanäle als durch Veränderungen des Arbeitsmarktstatus beeinflusst hat. So werden etwa die Stärkung des Selbstvertrauens und eine verbesserte soziale Inklusion als derartige Wirkungskanäle identifiziert. Zukünftige Wirkungsforschung auf dem Gebiet der aktiven Arbeitsmarktpolitik sollte daher derartige Kanäle mit in Betracht ziehen. Es erscheint in jedem Fall sinnvoll, das individuelle Wohlbefinden bereits in Datenerhebungen zu berücksichtigen.

Die Untersuchungen von lebenslangem Lernen in Form von berufsbegleitenden Weiterbildungsaktivitäten in Kapitel 5 ergeben zwar ein relativ unklares Bild hinsichtlich der Effekte auf die individuelle Lohnentwicklung, es zeigt sich jedoch ein positiver Effekt der Teilnahme auf die Wahrscheinlichkeit einer zukünftigen Erwerbstätigkeit. Dieser Effekt ist bis etwa fünf Jahre nach dem Ende der Weiterbildungsaktivitäten zu beobachten. Die Befunde legen außerdem den Schluss nahe, dass der Effekt in erster Linie auf die Tatsache einer Teilnahme zurückzuführen ist, während die jeweilige Dauer der berufsbegleitenden Weiterbildung keinen signifikanten Einfluss ausübt. Während dieses Ergebnis in Einklang mit der Signaltheorie von Humankapitalinvestitionen steht, ist es mit einem konkurrierenden Ansatz, in dem Humankapitalakkumulation die produktiven Fähigkeiten steigert, nur schwer vereinbar.

# Curriculum Vitae
# Ulf Rinne

*Aus datenschutzrechtlichen Gründen*
*wird in der elektronischen Version*
*auf einen Lebenslauf verzichtet.*

*Aus datenschutzrechtlichen Gründen
wird in der elektronischen Version
auf einen Lebenslauf verzichtet.*

*Aus datenschutzrechtlichen Gründen*
*wird in der elektronischen Version*
*auf einen Lebenslauf verzichtet.*

**Erklärung gem. § 9 der Promotionsordnung zum Dr. rer. pol. des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin vom 16. Juli 2008**

Hiermit versichere ich,
dass ich die vorliegende Dissertation selbstständig verfasst habe.

Bonn, im Februar 2009

*Ulf Rinne*