Essays on
Long–Run Employment Effects of
Training Programs for the Unemployed,
Regional Mobility,
and Data Quality

Inauguraldissertation zur Erlangung des akademischen Grades
eines Doktors der Wirtschaftswissenschaften
der Universität Mannheim

Robert Völter

vorgelegt im April 2008

Referent:	Prof. Bernd Fitzenberger, Ph.D.
Koreferent:	PD Dr. Friedhelm Pfeiffer
Abteilungssprecher:	Prof. Dr. Enno Mammen
Tag der mündlichen Prüfung:	15. September 2008

Acknowledgements

This dissertation could not have been written without the support of many people. First and foremost, I would like to thank my supervisor Bernd Fitzenberger. I am deeply indebted to him for his invaluable ideas and suggestions, his encouraging support, his guidance, and his patience. I am grateful to Friedhelm Pfeiffer for being the second supervisor of my dissertation.

My thesis work benefited greatly from cooperation with Bernd Fitzenberger and Aderonke Osikominu. Our joint research projects gave me the opportunity to learn so much about empirical labor economics.

I thank the participants in the doctoral seminars in Mannheim, Frankfurt and Freiburg for motivating discussions. In particular, I would like to mention my colleagues Annette Bergemann, Martin Biewen, Alfred Garloff, Karsten Kohn, Aderonke Osikominu, Stefan Speckesser, Alexandra Spitz-Oener, Marie Waller, and Ralf Wilke. The productive as well as pleasant, familial atmosphere at Bernd Fitzenberger's chair, first at the University of Mannheim and then at Goethe-University Frankfurt, certainly improved the quality of this thesis.

The Center for Doctoral Studies in Economics (CDSE) at the University of Mannheim provided me with sound graduate education and offered a stimulating academic environment. I gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG).

Last but not least, I would like to thank my parents for their continuous loving support. I dedicate this thesis to them.

Contents

General Introduction

The German labor market has been characterized by high unemployment rates for many years. In West Germany the unemployment rates stood at around 9% during most of the 1980s. In the 1990s they peaked at 11%. In East Germany the unemployment rates are in general around twice as high as in West Germany since the reunification. In the 1990s they rose to about 20%. Moreover, the unemployment rates in Germany show large regional variation. In 1992, for instance, in the West German federal states they ranged between 4.4% in Baden-Württemberg and 10.7% in Bremen. These unemployment rate differentials are very persistent.

One response from the government to the high unemployment rates has been the massive implementation of active labor market policy. Between 1990 and 1999 the German Federal Employment Agency has spent 208 billion Euro on active labor market policy. Training programs for the unemployed are the main element of active labor market policy. Therefore, it is essential to evaluate the effectiveness of such programs in raising the employment rates of their participants.

One response from individuals to large and persistent unemployment rate differentials can be to move from high to low unemployment regions.

In this dissertation the above outlined topics are studied using the IAB employment subsample (IABS). In chapter 1, long–run employment effects of training programs for the unemployed in West Germany are analyzed. In chapter 2, long–run effects under the considerably more difficult economic conditions in East Germany are evaluated. The topic of chapter 3 is the sensitivity of regional migration to regional unemployment rates in West Germany.

Chapter 4 concerns the data used. In empirical research data quality is a key factor influencing the credibility of the results. The data employed in this thesis, the IABS, is based on very reliable administrative data, the employment register of the German social insurance system. However, one of its variables of great interest for research, the education variable, suffers from quality deficiencies, namely missing values and wrong values. In chapter 4, deductive imputation procedures are proposed in order to improve the quality of the education variable in the IABS.

Each of the four chapters comprising this thesis has the format of a stand-alone paper. A brief overview is given as follows.

Chapter 1: Get Training or Wait? Long–Run Employment Effects of Training Programs for the Unemployed in West Germany
and
Chapter 2: Long–Run Effects of Training Programs for the Unemployed in East Germany

These two chapters are pooled and treated together for purposes of this introduction. This offers the possibility, on the one hand to avoid overlap, for the methods used are the same, and on the other hand to present in context the results obtained separately for West and East Germany. Chapter 1 is based on joint work[1] with Bernd Fitzenberger and Aderonke Osikominu, chapter 2 on joint work[2] with Bernd Fitzenberger.

In West Germany training programs for the unemployed have been used quite frequently compared to other countries. However, there has been considerable scepticism whether such large scale programs can be successful in raising the employment chances of their unemployed participants. In East Germany after reunification the same training programs have been used even more frequently than in the West and in a more challenging environment. The economy was in transition from a planned economy to a market one. The unemployment rates were very high. The administration was inexperienced with training programs resulting in very loosely targeted participation. Chapters 1 and 2 evaluate the differential employment effects of three typical training programs for the unemployed in West Germany and East Germany respectively.

The focus is on *long–run* employment effects. It is well known that in the short run training programs for the unemployed have negative employment effects, called lock–in effects. They are explainable because the effects are measured from program start onwards, and time in the programs counts as time out of employment. In the medium run some studies find positive employment effects. For the assessment of the programs it is crucial to know if any positive medium-run employment effects are really sustainable and persist in the long run. The contribution of these two chapters is to evaluate the effects for at least six years after program start.

---

[1] 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.

[2] Fitzenberger, B. and R. Völter (2007). "Long–Run Effects of Training Programs for the Unemployed in East Germany." *Labour Economics* 14(4), 730-755.

The objective of this analysis is to estimate the causal effect of participation in a training program on its participants. This effect is defined as the difference between the actual employment outcome of the participants and the counter-factual employment outcome they would have in case they would not have participated in the program. Since the counter-factual non-participation outcome of participants is unknown, one takes the known outcome of non-participants similar to the participants and uses the difference in outcomes between participants and similar non-participants as measure of the program effect.

To evaluate the programs two types of employment effects are analyzed. The first is the average effect on the employment rate (the share of participants in employment) in percentage points measured 4–6 years after program start. This employment rate effect excludes the years immediately after program start to focus on the long–run effects and to abstract from negative short–run lock–in effects. The second effect is the effect on the total time in employment measured in months during six years after program start. This total employment effect contrasts negative lock–in effects and any later positive effects. A positive total effect means that time lost during participation can be compensated by higher employment rates after participation.

The program effects are evaluated conditional on the unemployment duration before program start. Three stratums are considered, namely program starts during months 1–6 (stratum 1), months 7–12 (stratum 2) and months 13–24 (stratum 3) of the unemployment spell. Each stratum has its own control group of non-participants, who do not participate in any program during the stratum of unemployment considered. These non-participants can, however, include unemployed who participate in a program after the stratum considered. Such a dynamic definition of non-participants has the advantage that one evaluates the effect of the participation decision as it is actually made at the employment agency. Each time an unemployed meets his case worker the decision to be made is whether the unemployed starts a program *at this time* or not. The alternative, not to participate at this time, leaves open the option of participation at a later time in case the unemployed stays unemployed. The first paper's title "Get training or wait?" refers to such decision.

In a non-experimental evaluation, like the present one, the simple difference in outcomes between participants and non-participants is a biased measure of program effects. It reflects program effects as well as effects of differences in other characteristics, for example age. Hence, using the actual group of non-participants, a synthetic group of non-participants is constructed with the same average characteristics as the participants. These are the matched non-participants. Then the program effect is measured by the outcome difference between participants and matched non-participants. The method applied is local linear matching based on the propensity score, relying on the conditional independence assumption. The covariates used in the propensity score estimations include socio-economic characteristics, previous employment history, beginning of unemployment and elapsed duration of unemployment. The standard errors of the estimated average treatment effects on the treated are obtained from bootstrapping.

The three main training programs at that time are evaluated:

- SPST, provision of specific professional skills and techniques, is the most freqent program. It is taken by more than 60% of program participants. SPST is oriented towards theoretical skills imparted mainly in classroom courses, e.g. computer courses. Its duration is about six months.
- PF, practice firms, provide mainly practical skills in simulated training firms and workshops. The program PF also lasts about six months.
- RT, retraining, is a very comprehensive program and trains practical as well as theoretical skills. It results in a vocational training degree. RT takes 12–18 months, which is much longer than SPST and PF.

The evaluation is carried out using four inflow samples from employment into unemployment. These are in the respective inflow years: (1) West German women and men pooled in 86/87, (2) West German women and men pooled in 93/94, (3) East German women in 93/94 and (4) East German men in 93/94. The evaluations are always conducted separately for the three stratums of unemployment duration before program start. The share of unemployed participating in a training program in the western and eastern samples is around 10% and 20% respectivly. This demonstrates the more extensive use of training programs in the East.

Regarding results for West Germany, an overview of the long–run employment effects in the 93/94 cohort is given in table 1. The first outcome measure is the average effect on the employment rate. The results indicate that SPST increases the long–run employment rate of its participants by 13–16 percentage points. The long–run employment rate effects of PF and RT often reach this magnitude but in some cases are smaller. The second outcome measure is the effect on the total time spent in

Table 1: Long–Run Employment Effects West Germany 93/94

	SPST	PF	RT
Employment Rate Effects	13–16	0–16	7–18
Total Employment Effects	7–9	0–9	-3–4

Notes: Average effects on employment rate in percentage points during years 4–6 after program start. Total effects on time in employment measured in months during 6 years after program start. The ranges cover results for the three stratums. 0 denotes statistically insignificant effects.

employment in months during the six year period since program start. It is found that participants in SPST have a total employment gain of 7–9 months. The total employment effects of PF and RT only partly reach this magnitude. They are mostly smaller and in one case even negative. The effects in the 86/87 cohort in West Germany are in general comparable to the effects in the 93/94 cohort.

The conclusions are twofold. First, in general all three programs show sustainable and persistently positive long–run employment rate effects for at least six years. Second, based on rate and total effects, the most frequent program, SPST, delivers the best results for its participants. RT is no more effective than SPST or PF, despite providing 12–18 months of training instead of only about 6 months.

Given SPST is successful for its own participants, a natural further question is if participants in PF or RT would have gained by participating in SPST instead. Hence, in a multiple evaluation, pairwise comparisons between the effects of participation in any two of the three programs are performed. In the long run, most differences in employment effects between two programs prove to be insignificant due to large standard errors. So, little evidence is found that a reallocation of participants between the programs would have been beneficial for the employment outcomes.

Regarding East Germany, long–run employment effects are shown in table 2a. SPST is discussed first. In the East SPST shows positive long–run average effects on the employment rate of 7–13 percentage points. It can be concluded that in East Germany also SPST causes positive and persistent long–run employment rate effects. However, under the difficult circumstances of high unemployment and little targeting its employment rate effects are about a third lower than in the West. The total employment gains from participation in SPST lie between 0 and 4 months, considerably lower than in West Germany. This is explained by longer program durations and much stronger negative lock–in effects in the East.

Table 2a: Long–Run Employment Effects East Germany 93/94

	SPST Woman/Men	PF Woman/Men	RT Women/Men
Employment Rate Effects	7–10/7–13	0-0/0	0–16/0–12
Total Employment Effects	0–3/0–4	0-0/0	0-0/-5–0

Table 2b: Long–Run Effects on Unemployment Benefit Recipiency
East Germany 93/94

	SPST Woman/Men	PF Woman/Men	RT Women/Men
Unempl. Ben. Rate Effects	0-0/0-0	0–12/0	0-0/0–13
Total Unempl. Ben. Effects	0–7/0–9	0–11/7	8–12/8–14

Notes: Average effects on 2a: employment rate (2b: unemployment benefit recipiency rate) in percentage points during years 3–6 after program start. Total effects on time in 2a: employment (2b: unemployment benefit recipiency) measured in months during 6 years after program start. The ranges cover results for the three stratums. 0 denotes statistically insignificant effects.

In contrast to SPST, the conclusions about long–run employment effects of PF and RT in East Germany are rather negative. RT only occasionally increases the long–run employment rate. It never increases the total time in employment. PF does not show any positive employment effects.

Individuals can be in three labor market states: (i) employment, (ii) unemployment or (iii) non-employment, i.e. having left the labor force. Higher employment rates can be accounted for by lower unemployment rates or lower non-employment rates. In order to differentiate between the two, the long–run average effects on the unemployment benefit recipiency rate in East Germany are evaluated using the same methodology as in evaluating employment effects. Table 2b shows that no program results in a reduction in the rate of unemployment benefit recipients. Mostly this rate is unchanged in the long run, sometimes the rate of unemployment benefit recipients even increases. This suggests the conclusion that successful programs raise the employment rate of their participants by preventing them from leaving the labor force. Table 2b also shows that in East Germany program participation increases the total time receiving unemployment benefits. The fact that participants receive unemployment benefits while being in a program explains this mostly.

Finally, regarding gender differences, the long–run effects of participation in a training program in East Germany are found to be quite similar for women and men.

Chapter 3: Regional Unemployment and Regional Mobility in West Germany

The topic of chapter 3 is the link between the regional unemployment rate and regional mobility. The large and persistent geographical unemployment rate differentials in Germany raise the question if individuals respond by moving from high to low unemployment regions. Such migration could help equalize unemployment rate differentials. From a macroeconomic perspective this can be beneficial for the economy by enhancing efficiency. Taking the microeconomic perspective, migration to low unemployment regions can also be beneficial, for unemployed as well as for employed individuals. It is an opportunity to increase medium– to long–run employment prospects, because it decreases the risk of becoming unemployed in the future. This is true as long as unemployment rate differentials are persistent as in Germany. For the unemployed migration to lower unemployment regions can also be a way to get a job in the short run.

The investigation is based on an analysis of job changes. A sample of job changes of males in West Germany between 1984 and 1997 is obtained using the IABS. The measure of regional mobility is the share of job changes between regions among all job changes. A job change between regions is defined as job change in which the job changer leaves his region and takes a new job at least 100 km away. Measured in this manner, mobility stands at around 9 percent in the sample considered.

The main question examined is how the share of regional mobility varies with the regional unemployment rate. It is analyzed in a linear probability model controlling for individual characteristics of the job changers. It is found that regional mobility is sensitive to regional unemployment. A one percentage point higher regional unemployment rate is linked to a 0.1–0.2 percentage points higher share of job changers, which leave the region. However, a 0.1–0.2 percentage points increase per percentage point increase in the regional unemployment rate is small compared to the average share of regional mobility of around 9 percent. It is also small compared to the influence of individual factors like employment status or education. University graduates exhibit an about 8 percentage points higher share of regional mobility than men with a vocational training degree.

This chapter's main finding is a statistically significant but economically very weak positive relationship between the regional unemployment rate and regional mobility. This result is consistent with the observed high persistence of regional unemployment rate differentials.

It is also investigated whether the relationship between the regional unemployment rate and regional mobility is heterogeneous. Regarding age it is found, that only for those up to the age of 40 regional mobility is sensitive to regional unemployment. Differentiating the decades under investigation, the 80s and the 90s, the sensitivity is only significant in the 80s.

An additional aspect is examined in this chapter. Looking at job changes between regions, i.e. the 9% of all job changes covering distances of at least 100 km, the average difference in the regional unemployment rates between origin and destination regions is analyzed. It is found that in the destination regions the unemployment rate is 0.3 percentage points lower in the 80s and 0.1 percentage points lower in the 90s respectively. This shows the tendency of the regionally mobile individuals, to take new jobs in regions with lower unemployment rates.

Chapter 4: Imputation Rules to Improve the Education Variable in the IAB Employment Subsample

Chapter 4 is based on joint work[3] with Bernd Fitzenberger and Aderonke Osikominu.

This thesis is build on the IAB employment subsample (IABS) as data source. The IABS is a one percent employment sample drawn from the employment register data of the German social insurance system. The main advantages of the IABS compared to survey data are its large size and the reliability of core information like those on earnings. However, some information like the one on formal education is less reliable. The likely reason is that this information is not necessary for the original purpose of the data, the social insurance administration.

For many issues studied in labor economics a reliable measure of education is important. Two leading examples are returns to education and skill biased technological change. The education variable given in the IABS data shows two deficiencies. First, there are missing values for about 10% of the spells. And second, there are incorrectly reported educational degrees resulting in misclassifications. This concerns some spells of up to 20% of the individuals in the data set.

[3]Fitzenberger, B., A. Osikominu and R. Völter (2006). "Imputation Rules to Improve the Education Variable in the IAB Employment Subsample." *Schmollers Jahrbuch (Journal of Applied Social Science Studies* 126, 405-436.

The technical chapter 4 addresses these deficiencies and proposes three deductive, nonstochastic imputation procedures to improve the education variable in the IABS. They are based on specific characteristics of formal education. Formal education is mostly constant once an individual has entered working life. In case it changes it can only increase, by attainment of an additional degree. It never decreases. Finally, the employers have to report the highest degree an individual has attained. These features in combination with the panel structure of the data make it possible to improve the education variable by extrapolating degrees. Degrees deemed reliable are extrapolated to spells with missing education and to spells which are deemed unreliable. The critical decision is which degrees can serve as basis for extrapolation.

The proposed imputation procedures differ in the heuristic rules, which reports are to extrapolate. The procedures are designed to bound the true education in distribution from above and below. One imputation procedure extrapolates almost every report, including likely overreports, and thus can be viewed as giving an upper bound for the true education. Two other procedures are designed as more conservative procedures, resulting in lower bounds for the true education level. The second procedure does not extrapolate rarely reported degrees. The third procedure does not extrapolate reports coming from employers with bad reporting quality.

The results of the imputation procedures are as follows. Imputation removes more than two thirds of the missing values. Concerning the education distribution of employment, the improved data show the educational attainment of the labor force to be higher than the raw data suggest. All procedures display higher shares of holders of vocational training, technical college and university degrees and lower shares of individuals without any of these degrees. The resulting shares in general do not differ essentially in size between the procedures. Regarding wage inequality, improving the data affects especially some measures of wage inequality at lower percentiles in the low and medium skilled groups. In Mincer-type wage regressions, improving the data makes only a small difference. However, wages for individual skill groups are typically lower at all degrees when using the imputed data.

It can be concluded that in order to ensure data quality, applying the easily implementable imputation procedures to improve the education variable in the IABS is a relevant step in preparing the data for empirical research. It changes the results in some typical applications. In this thesis the improved education variable is used in chapters 1 and 2 in the matching procedures and in chapter 3 in the regressions.

Chapter 1

Get Training or Wait? Long–Run Employment Effects of Training Programs for the Unemployed in West Germany

1.1 Introduction

Public sector sponsored training has traditionally been a main part of active labor market policy (ALMP) in many countries like Germany. During the last decade, there were many pessimistic assessments regarding the usefulness of public sector sponsored training programs in raising employment chances of the unemployed (see the surveys in Fay (1996), Heckman et al. (1999), Martin and Grubb (2001), Kluve and Schmidt (2002)). While the surveys emphasized that small scale training programs, which are well targeted to specific groups and which involve a strong on–the–job component, can show positive employment effects, these studies doubt that the large scale training programs in countries like Germany are successful in raising on average the employment chances of adult workers who became unemployed and who participate in such programs. Negative short–run effects of these programs are attributed to the lock–in effect while being in the program.

Recently, OECD (2005) has emphasized that long–term labor market programs, such as training, often have little or negative short–run effects on outcomes. Also, it is clear that lock–in effects are worse for longer programs, because they keep the unemployed away from the labor market for a longer time. However, it could be the case that sizeable labor market effects are only to be expected from sufficiently long training programs (Fay, 1996). Therefore, it is crucial to assess program impacts in a longer term perspective in order to investigate whether the sizeable lock–in effects in the short run are compensated by positive long–run effects. In fact, OECD (2005) reports positive long–term results for some training programs. Our paper adds to this literature by estimating the long–run employment effects of three different types of training programs in Germany over at least six years since the beginning of the treatment.

The vast majority of the evaluation studies summarized in the aforementioned surveys used a static evaluation approach receiving treatment during a certain period of time against the alternative of not receiving treatment during this period of time. In a dynamic setting, the timing of events becomes important, see Abbring and van den Berg (2003, 2005), Fredriksson and Johansson (2003, 2004), and Sianesi (2003, 2004). Static treatment evaluations implicitly condition on future outcomes leading to possibly biased treatment effects. This is because the nontreated individuals in the data might be observed as nontreated because their treatment starts after the end of the observation period or because they exit unemployment before treatment starts (Fredriksson and Johansson 2003, 2004). This paper follows Sianesi (2003, 2004) and estimates the effects of treatment starting after some unemployment experience against the alternative of not starting treatment at this point of time and waiting longer.

For Germany, appropriate data for a long–term evaluation of public sector sponsored training were not available for a long time and there existed serious scepticism in the German policy debate as to whether ALMP is actually effective (Hagen and Steiner, 2000). Until recently, basically all the evaluation studies[1] made use of survey data.[2] Although these data are rich with respect to informative covariates, the evaluation studies using survey data suffer from severe shortcomings with respect to the quality of the treatment information and the precision of the employment history before and after treatment. The sample sizes in these studies are typically small. They do not allow the researcher to evaluate the effects of any heterogeneous treatment or of treatments targeted to specific groups of individuals.

Contributing to the debate on the effectiveness of ALMP, this paper analyzes the employment effects of three types of public sector sponsored training programs in West Germany. We use unique administrative data which have only recently become available. Using data on employment, periods of transfer payments, and participation in training programs, we carefully identify three types of public sector sponsored training programs for the unemployed. These programs are not associated with a regular job. The largest program among the three is the Provision of Specific Professional Skills and Techniques (SPST). SPST programs provide additional skills and specific professional knowledge in courses with a median duration between 4 and 6 months. The two other training programs are working in a Practice Firm (PF) and Retraining (RT). RT involves a long program, which lasts up to 2 years and which provides complete vocational training in a new occupation. PF involves training in a work environment simulating a real job and is of similar length as SPST. This classification of treatments is developed in this paper and in the earlier

[1] See Speckesser (2004, chapter 1) and Wunsch (2006, section 6.5) as recent surveys for Germany. Previous studies based on survey data gave inconclusive evidence. For instance, for East Germany, Lechner (2000) found negative employment effects of training programs in the short run and insignificant effects in the long run based on survey data. In contrast, Fitzenberger and Prey (2000) found some positive employment effects of training programs in East Germany.

[2] Notable exceptions are the recent studies of Lechner et al. (2005a,b) and Fitzenberger and Speckesser (2007), which are all based on the same data set as our study. In fact, the data set is the outcome of a joint effort to merge administrative data for evaluation purposes, see Bender et al. (2005).

paper, Fitzenberger and Speckesser (2007). The three training programs considered here differ both in length and content. PF has the strongest on–the–job component, SPST involves typically off–the–job classroom training and RT involves both on–the–job and off–the–job training for a specific occupation. Based on the aforementioned evidence reported by Martin and Grubb (2001) and others, PF should be the most effective program, at least in the short run. In contrast, Lechner et al. (2005a) report quite favorable evidence for RT.

This paper takes advantage of unique administrative data which integrate register data on employment as well as data on unemployment and participation in active labor market programs generated by the Federal Employment Office (*Bundesanstalt für Arbeit*, BA). Our data set merges register data with benefit data and with survey data obtained from the local offices of the Federal Employment Office. This survey records all cases of participation in further training programs during the period 1980–1997 and offers rich information on heterogeneous courses. Our analysis evaluates the effects of training for inflows into unemployment for the years 1986/87 and 1993/94 in West Germany. These two inflow samples faced very different labor market prospects due to changing business cycle conditions and the impact of German unification. It is of interest to investigate whether the effects of ALMP differ by the state of the labor market. The 1986/87 sample faced a fairly favorable labor market in the years to come culminating in the unification boom in West Germany. In contrast, the 1993/94 sample entered unemployment during one of the most severe recessions in West Germany resulting in a prolonged period with bad labor market chances.

Since our analysis is based on administrative data, we have to use a non–experimental evaluation approach. We build on the conditional independence assumption which says that for treated and nontreated individuals the expected potential employment outcome in case of not receiving the treatment of interest (which is observed for the nontreated and counterfactual for the treated) is the same conditional on a set of observable covariates. In our case, these covariates involve socio–economic characteristics, previous employment history, beginning of unemployment, and elapsed duration of unemployment. The analysis uses the popular propensity score matching approach adjusted to a dynamic setting building on the recent work by Fredriksson and Johansson (2003, 2004) and Sianesi (2003, 2004). In fact, when the timing of treatment is a random variable depending upon elapsed duration of unemployment, a static evaluation approach does not seem appropriate. We evaluate the employment effects of three multiple exclusive training programs both against the alternative of nonparticipation and in pairwise comparisons building on Lechner (2001) and Imbens (2000). Our matching estimator is implemented using local linear matching (Heckman, Ichimura, Smith and Todd 1998) based on the estimated propensity score. In fact, kernel matching has a number of advantages compared to nearest neighbor matching (Heckman, Ichimura, Smith and Todd (1998), Ichimura and Linton (2001), Abadie and Imbens (2006)), which is widely used in the literature (e.g. Lechner et al. (2005a,b), Sianesi (2003, 2004)). We run separate analyses conditional on elapsed duration of unemployment at the beginning of treatment. We distinguish between training programs starting during quarters 1 to 2, 3 to 4, and 5 to 8 of unemployment.

Our analysis extends considerably upon the earlier work of Fitzenberger and Speckesser (2007) in several dimensions. The earlier paper evaluates the employment effects of SPST against the comprehensive alternative of nonparticipation in SPST for 36 months after the beginning of the treatment. The analysis is performed only for the 1993 inflow sample into unemployment, both for East and West Germany. This study analyzes the effects of three exclusive training programs for inflow samples in 1986/87 and 1993/94 in West Germany. The three programs are analyzed in a multiple treatment framework and we evaluate medium– and long–run treatment effects up to 25–31 quarters after the beginning of the treatment depending on the start date of the treatment.

Comparing our work to the study by Lechner et al. (2005a), who also estimate multiple treatment effects based on the same data, there are the following notable differences. First, Lechner et al. only consider a sample for the 1990s, while we also consider an inflow sample from the 1980s. Second, their study uses a different approach regarding the construction of the sample and the choice of valid observations. The definition of treatment types and the identification of treatments from the data differ as well. Third, taking into account the dynamic assignment into programs our paper also comprises an important methodological difference compared to Lechner et al. (2005a) who apply a static approach. Nonetheless, as far as comparable, the results are quite similar in most cases.

The remainder of this paper is structured as follows: Section 1.2 gives a short description of the institutional regulation and participation figures for Active Labor Market Policy. Section 1.3 focuses on the different options of further training, their target groups, and course contents. Section 1.4 describes the methodological approach to estimate the treatment effects. The empirical results are discussed in section 1.5. Section 1.6 concludes. The final appendix provides further information on the data and detailed empirical results. An additional appendix includes further details on the data and on the empirical results.

1.2 Basic Regulation of Further Training

1.2.1 Programs

For the period covered by our data, further training in Germany is regulated on the basis of the Labor Promotion Act (*Arbeitsförderungsgesetz*, AFG) and is offered and coordinated by the German Federal Employment Office (formerly *Bundesanstalt für Arbeit*, BA). Originally, further training was conceived to improve occupational flexibility and career advancement and to prevent skill shortages. In response to unemployment becoming an increasingly persistent phenomenon during the 1980s and 1990s, further training changes its character from a rather preventive ALMP towards an intervention policy predominantly targeted to unemployed and to those at severe risk of becoming unemployed. With the increasing number of unemployed entering training programs, skill–upgrading courses targeted to employed workers lose importance in favor of courses in which individuals are taught new technologies or are given the opportunity to enhance existing skills for the purpose of occupational reintegration.

The German legislation distinguishes three main types of training: further vocational training, retraining, and integration subsidy. In addition, there is short–term training which only existed from 1979 to 1992 (§41a AFG). Although, during the 1980s and 1990s, there have been many changes concerning passive labor market policy – i.e. changes in benefit levels and eligibility criteria – the regulation of the traditional training schemes, further vocational training, retraining, and integration subsidy, remained stable until the end of 1997 when the Labor Promotion Act was replaced by the Social Code III. In the following, we give a short description of the main programs:[3]

[3]The complete list with descriptions of the different training schemes that are regulated by the Labor Promotion Act can be found in the additional appendix.

- **Further vocational training** (*Berufliche Fortbildung*) includes the assessment, maintenance and extension of skills, including technical development and career advancement. The duration of the courses depends on the participants' individual predispositions, other co–financing institutions and on the supply of adequate training courses by the training providers.

- **Retraining** (*Umschulung*) enables vocational re–orientation if a completed vocational training does not lead to adequate employment. Retraining is supported for a period up to 2 years and aims at providing a new certified vocational education degree.

- The program **integration subsidy** (*Einarbeitungszuschuss*) offers financial aid to employers who are willing to give employment to unemployed or to workers directly threatened by unemployment. The subsidy (up to 50% of the standard wage in the respective occupation) is paid for an adjustment period until the supported person reaches full proficiency in the job.

- In 1979, **short–term training** was introduced under §41a AFG aiming at "increasing prospects of integration". With this program, skill assessment, orientation and guidance should be offered to unemployed. The curricula under this program are usually short–term, lasting from two weeks up to two months and are intended to increase the placement rate of the unemployed. This type of training was abolished in 1992.

1.2.2 Financial Incentives for Participation

Except for the integration subsidy which is a subsidy to a standard salary (according to union wage contracts), participants in training programs are granted an income maintenance (IM, *Unterhaltsgeld*) if they satisfy the conditions of entitlement. To qualify, they must meet a minimum work requirement of being previously employed during at least one year in a job subject to social insurance contributions or they must be entitled to unemployment benefits or subsequent unemployment assistance.[4]

Starting 1986 until 1993, the income maintenance amounted to 73% of the relevant previous net earnings for participants with at least one dependent child and 65%

[4] If a person does not fulfill the requirement of previous employment, but had received unemployment assistance until the start of the program, an income maintenance may be paid as well.

otherwise. This was higher than the standard unemployment benefits (UB, *Arbeitslosengeld*) in this period which was at 68% and 63%, respectively. And IM was considerably higher for those unemployed whose UB expired and who were receiving the lower, means tested unemployment assistance (UA, *Arbeitslosenhilfe*) which amounted to 58% (with children) and 56% (without children).[5] In 1994, income maintenance and unemployment benefits were both cut back to a common level of 67% (with children) and 60% (without children), reducing the financial incentives to join a training program. Unemployment assistance was also lowered to 57% (with children) and 53% (without children).[6]

IM recipiency during a training program did not affect the entitlement period for unemployment benefit payments. Effectively, this means that the unemployed could defer the transition from unemployment benefits to unemployment assistance by taking part in a training program. Additionally, participants in training programs could requalify for unemployment benefits providing additional incentives to participate.[7]

Summing up, for our time period under investigation, there are positive financial incentives for the unemployed to join a program. The income maintenance is at least as high or higher than unemployment benefits and it is always higher than unemployment assistance. Furthermore, participation allows to postpone the transition from unemployment benefits to the lower, means tested unemployment assistance and sometimes even allows to requalify for unemployment benefits. In addition, the BA bears all costs directly incurred through participation in a further training scheme, especially course fees.

1.2.3 Participation

Participation in further training programs in West Germany is large, see table 1.2. In 1980, 247,000 participants enter such programs. In the late 1980s and early 1990s, annual entries peak at almost 600,000 and then decline to 378,400 in 1996. Among

[5] In the relevant period the exhaustion of UB and transition from the higher UB to the lower UA took place between the 6th and the 32nd month of unemployment, depending on age and employment history (for details see Plassmann, 2002).

[6] For detailed descriptions of the changes in regulations over time see Bender et al. (2005) and Steffen (2005).

[7] This is because, until 1997, periods of income maintenance payments were counted on the minimum work requirement for receiving unemployment benefits, for details see Bender and Klose (2000).

the three main types of training programs distinguished by German legislation, the general further vocational training schemes traditionally are the most important in West Germany with about 70–80% of the entries. Roughly 20% enter retraining. The small remaining share are integration subsidies.

1.3 Data and Treatment Types

We use a database which integrates administrative individual data from three different sources (see Bender et al. (2005) for a detailed description). The data contain spells on employment subject to social insurance contributions, on transfer payments by the Federal Employment Office during unemployment, and on participation in further training schemes.

The core data for this evaluation are taken from the **IAB Employment Subsample** (*IAB Beschäftigtenstichprobe*, IABS) of the Institute for Employment Research (IAB), see Bender et al. (2000) and Bender et al. (2005, chapter 2.1). The IABS is a 1% random sample drawn from employment register data for all employees who are covered by the social security system over the period 1975–97. Because sampling is based on employment, we restrict the analysis to inflows from employment to unemployment.[8] For this study, we obtained additional information from the IAB for 1998–2002, which we merged to the basic data.

The second important source is the **Benefit Payment Register** (*Leistungsempfängerdatei*, LED) of the Federal Employment Office (BA), see Bender et al. (2005, chapter 2.2). These data consist of spells on periods of transfer payments granted to unemployed and program participants from the BA. Besides unemployment benefit or assistance, these data also record very detailed information about income maintenance payments related to the participation in further training schemes.

The third data source are the data on training participation (FuU-data). The employment agency collects these data for all participants in further training, retraining, and other training programs for internal monitoring and statistical purposes,

[8]Restricting the analysis to inflows from employment, we exclude program participants who have not been employed before registering as unemployed. This restriction of the data only concerns a small share of training participants. Statistics from the BA show that the share of those who have never worked or have not worked within the last 6 years before they enter a program is between 4 and 8 percent. Because of the sample design, it would be impossible to construct an appropriate control group for such participants.

see Bender et al. (2005, chapter 2.3). For every participant, the FuU-data contains detailed information about the program and about the participant.

The FuU–data were merged with the combined IABS–LED data by social insurance number and additional covariates. Numerous corrections were implemented in order to improve the quality of the data, see Bender et al. (2005, chapters 3–4) and the additional appendix for details. The IABS provides information on personal characteristics and employment histories. The combination of the transfer payment information and the participation information is used to identify the likely participation status regarding the different types of training programs.

1.3.1 Evaluated Programs

We evaluate the three quantitatively important training programs targeted at the unemployed. Further vocational training is a very broad legal category and consists of quite heterogeneous programs. Hence we utilize a classification developed in Fitzenberger and Speckesser (2007) and evaluate two specific further vocational training programs: Practice Firms (PF) and provision of specific professional skills and techniques (SPST). We also evaluate Retraining (RT). We do not evaluate short–term training, since this program only existed until 1992.[9] We also do not evaluate integration subsidies, because they condition on having found a job involving a potentially difficult endogeneity issue.[10] Next, we describe the evaluated programs. Table B.24 gives an overview of the evaluated programs.

Practice Firms are simulated firms in which participants practice everyday working activities. The areas of practice are whole fields of profession, not specific professions. Hence, practice firms mainly train general skills while the provision of new professional skills is less important. Some of the practice firms are technically oriented, the practice studios, whereas others are commercially oriented, the practice enterprises. One of the practice firms' goals is to evaluate the participant's aptitude for a field of profession. The programs usually last for six months and do not provide official certificates.

[9] After 1992 comparable programs were offered, but they are not recorded in the data. In order to analyze both inflow samples in a comparable way we ignored information about short–term training in the eighties.

[10] A detailed description of how the different treatment types are identified from the data is given in the additional appendix.

Provision of **Specific Professional Skills and Techniques** intends to improve the starting position for finding a new job by providing additional skills and specific professional knowledge in medium–term courses. It involves refreshing specific skills, e.g. computer skills, or training on new operational practices. SPST mainly consists of classroom training but an acquisition of professional knowledge through practical work experience may also be provided. After successfully completing the course, participants usually obtain a certificate indicating the contents of the course, i.e. the refreshed or newly acquired skills and the amount of theory and practical work experience. Such a certificate is supposed to serve as an additional signal to potential employers and to increase the matching probability since the provision of up–to–date skills and techniques is considered to be a strong signal in the search process. The provision of specific professional skills and techniques aims at sustained reintegration into the labor market by improving skills as well as providing signals.

Compared to retraining, which is a far more formal and thorough training on a range of professional skills and which provides a complete vocational training degree, SPST provides a smaller, specific addition to the occupational knowledge. However, this addition certainly exceeds the level provided in short–term programs (not evaluated here) that usually aim at improving job search techniques or general social skills. Thus, SPST ranges in the middle between very formal (and very expensive) courses and very informal and short courses (improving general human capital).

Retraining consists of the provision of a new and comprehensive vocational training according to the regulation of the German apprenticeship system. It is targeted to individuals who had already completed a vocational training degree and face severe difficulties in finding a new job within their profession. It might however also be offered to individuals without a first formal training degree if they fulfill additional eligibility criteria.

Retraining provides widely accepted formal certificates. It comprises both, theoretical training and practical work experience. The theoretical part of the formation takes place in the public education system. The practical part is often carried out in firms that provide work experience in a specific field to the participants, but sometimes also in interplant training establishments. This type of treatment leads to a certified job qualification in order to improve the job match. Ideally, the training occupation in retraining corresponds to qualifications which are in high demand in the labor market.

Table 1.1: Overview of Evaluated Programs

	PF	SPST	RT
Name	Practice Firm	Provision of Specific Skills and Techniques	Retraining
Description	training on the job in a simulated firm	classroom courses	complete new vocational training
Orientation	practical	theoretical	practical and theoretical
Median duration 86/87 (93/94)	5 (6) months	4 (6) months	12 (16) months

1.3.2 Inflow Sample into Unemployment and Participation by Type of Training

The goal of this study is to analyze the effect of training programs on employment chances of unemployed individuals. Therefore, we base our empirical analysis on inflow samples into unemployment. We use the inflows into unemployment in the years 1986/87 and 1993/94 in West Germany, omitting Berlin and East Germany. Effectively, we consider individuals who experience a transition from employment to nonemployment and for whom a spell with transfer payments from the Federal Employment Office starts within the first twelve months of nonemployment or for whom the training data indicate a program participation before the unemployed individual finds a new job.[11] In the following, we denote the start of the nonemployment spell as the beginning of the unemployment spell. We condition on receipt of unemployment compensation or program participation to exclude most of the individuals who move out of labor force after exiting from their job. This concerns especially individuals whose treatment status would be nonparticipation in any training program during their nonemployment spell. A treatment is associated with an unemployment spell, if the individual starts training before possibly exiting to employment. In our monthly data, this means that the individual should still be recorded as nonemployed in the month when treatment starts. Furthermore, we restrict our samples to the 25 to 55 years old in order to rule out periods of formal education or vocational training as well as early retirement.

We choose the years 93/94 and 86/87 to allow for a comparison between the 1980s and the 1990s. Figure 1.1 depicts the unemployment rate in West Germany. The

[11]We allow the same individual to appear in the sample more than once if he or she exhibits more than one transition from employment to unemployment during the relevant time period.

dotted vertical lines mark the years 1986 and 1993, respectively. Whereas 86/87 mark the end of a sequence of years with relatively high unemployment, the cohort 93/94 enters during a period with increasing unemployment rates. Thus, the 86/87 cohort faced a fairly favorable labor market in the years to come culminating in the unification boom in West Germany, while the 93/94 cohort entered unemployment during one of the most severe recessions in West Germany resulting in a prolonged period with bad labor market chances. Our data allow to follow individuals entering unemployment in 86/87 until December 1996/97 and individuals entering unemployment in 93/94 until the end of 2001/02.

Table 1.3 gives information about the size of the inflow samples and the incidence of training. We focus on the three types of training programs which are most suitable for unemployed individuals and which do not involve on–the–job training (training while working in a regular job). These are (i) practice firm (PF), (ii) provision of specific professional skills and techniques (SPST), and (iii) retraining (RT). The total inflow sample comprises 20,902 spells for the 86/87 cohort and 25,051 spells for the 93/94 cohort. There are 1,714 training spells for the eighties and 2,727 for the nineties. Thus, about 10% of all unemployed participate in one of the three training programs considered. Among these, SPST represents by far the largest type of training with 64% and 72% of the training spells, respectively in the two samples. About one fifth of all training spells are RT, and PF represents the smallest group in both samples. In absolute numbers, there are 246 (325) PF spells in the 86/87 (93/94) inflow sample, 1,093 (1,944) SPST spells and 375 (458) RT spells. Table 1.4 shows the frequency of training by time window of elapsed unemployment.

Table 1.5 provides descriptive statistics on the elapsed duration of unemployment at the beginning of treatment. Our discussion focuses on quantiles because averages can be biased due to outliers. The median entrant in PF has been unemployed for 10 months in the 86/87 sample and 9 months in the 93/94 sample. Late starts (75%–quantile) of PF occur after 19 months in the 86/87 sample and much earlier in the 93/94 sample. For RT, the quantiles in the samples are very similar. With a median of 6 and 7 months, RT starts the earliest. For SPST, we find a reversed trend in comparison to PF. While SPST participation starts almost as early as RT in the 86/87 sample, the starting dates are noticeably later in 93/94, with the median increasing from 6 to 11 months.

Table 1.6 provides descriptive information on the realized duration of training spells.

The average duration of practice firm is similar in both samples with 5.1 months in 86/87 and 5.7 months in 93/94. SPST has an average duration of 4.9 months for the 86/87 sample and 6.3 months for 93/94. Retraining is by far the longest program. It lasts on average for 13.1 months in the 86/87 sample and 14.9 months in the 93/94 sample. Note that some participants drop out of the programs early. So the realized durations can be shorter than the planned durations. In our samples about 70% of the participants complete the programs, 10% drop out because they have found a job and 20% drop out for other reasons. Our analysis does not condition on program completion since program dropout is likely to be endogenous. So, strictly speaking, the programs we evaluate are the starts of the respective programs, as is common in most of the recent literature.

The final question about the samples which we want to discuss is the incidence of other programs. Our basic approach is to ignore the (relatively rare) participation in other programs and classify such spells as spells without program participation. For 86/87, 1.2% of the participants in evaluated programs participated in short–term training before starting an evaluated program. Also 1.2% of the nonparticipants took part in short–term training during their defining unemployment spell.[12] For 93/94, there existed no short–term training. The share of nonparticipants in the evaluated programs who took part in another, not evaluated further vocational training program[13] is 0.5% in the 86/87 sample and 0.3% in the 93/94 sample. Integration subsidies are paid to 1.2% of the nonparticipants in the 86/87 sample. Only 0.3% of the training participants in the 86/87 sample finish their unemployment spell with a subsidized job. And in the 93/94 sample among both, the treated and the controls, this share is 0.3% or lower. Concluding, we argue that participation in other, not evaluated training programs is small enough to be neglected in our analysis.

1.4 Evaluation Approach

Our goal is to analyze the effect of $K = 3$ different training programs on the quarterly employment rate at the individual level, which is measured as an average of

[12] As we do with the evaluated programs we look at program participation during the defining unemployment spell and in the case of integration subsidies at payment of a subsidy for the first job after the defining unemployment spell.

[13] These other programs are mainly career advancement programs targeted at the employed.

three monthly employment dummy variables.[14] In a situation where individuals have multiple treatment options, we estimate the average treatment effect on the treated (ATT) of one training program against nonparticipation in any of the three programs and of pairwise comparisons of two programs. Extending the static multiple treatment approach to a dynamic setting, we follow Sianesi (2003, 2004) and apply the standard static treatment approach recursively depending on the elapsed unemployment duration. The implementation builds upon the approach for binary treatment in Fitzenberger and Speckesser (2007). We first present our evaluation approach and then compare it to recent alternative proposals in the literature.

1.4.1 Multiple Treatments in a Dynamic Context

Our empirical analysis is based upon the potential–outcome–approach to causality, see Roy (1951), Rubin (1974), and the survey of Heckman et al. (1999). Lechner (2001) and Imbens (2000) extend this framework to allow for multiple, mutually exclusive treatments. Let the 4 potential outcomes be $\{Y^0, Y^1, Y^2, Y^3\}$, where $Y^k, k = 1, ..., 3$, represents the outcome associated with training program k and Y^0 is the outcome when participating in none of the 3 training programs. For each individual, only one of the $K + 1$ potential outcomes is observed and the remaining K outcomes are counterfactual. We estimate the average treatment effect on the treated (ATT) of participating in treatment $k = 1, 2, 3$ against nonparticipation $k = 0$ (treatment versus waiting) and the differential effects of the programs (program k versus program l where $k, l \neq 0$), see Lechner (2001).

Fredriksson and Johansson (2003, 2004) argue that a static evaluation analysis, which assigns unemployed individuals to a treatment group and a nontreatment group based on the treatment information observed in the data, yields biased treatment effects. This is because the definition of the control group conditions on future outcomes or future treatment. For Sweden, Sianesi (2004) argues that all unemployed individuals are potential future participants in active labor market programs, a view which is particularly plausible for countries with comprehensive systems of active labor market policies (like Germany). In former West Germany, active labor market programs were implemented at a fairly large scale in international comparison. This discussion implies that a purely static evaluation of the different training programs is not warranted. Following Sianesi (2003, 2004), we analyze the effects

[14]The quarterly employment rate can take the four values 0, 1/3, 2/3, and 1.

of the first participation in a training program during the unemployment spell considered *conditional on the starting date of the treatment*. We distinguish between treatment starting during quarters 1 to 2 of the unemployment spell (stratum 1), treatment starting during quarters 3 to 4 (stratum 2), and treatment starting during quarters 5 to 8 (stratum 3).

We analyze treatment conditional upon the unemployment spell lasting at least until the start of the treatment k and this being the first treatment during the unemployment spell considered. Therefore, the ATT parameter (comparing treatments k and l) of interest is

$$(1.1) \qquad \theta(k,l;u,\tau) = E(Y^k(u,\tau)|T_u = k, U \geq u-1, T_1 = ... = T_{u-1} = 0)$$

$$-E(Y^l(\tilde{u},\tau-(\tilde{u}-u))|T_u = k, u \leq \tilde{u} \leq \bar{u}, U \geq u-1, T_1 = ... = T_{u-1} = 0)\ ,$$

where T_u is the treatment variable for treatment starting in quarter u of unemployment. $Y^k(u,\tau)$, $Y^l(u,\tau)$ are the potential treatment outcomes for treatments k and l, respectively, in periods $u+\tau$, where treatment starts in period u and $\tau = 0,1,2,...$, counts the quarters since the beginning of treatment. When $l = 0$, we compare treatment k versus waiting (nonparticipation in the stratum) and when $l \geq 1$, we do a pairwise comparison between treatment k and l. U is the duration of unemployment, $\tilde{u}$ is the random quarter when alternative treatment l starts, and $\bar{u} = 2,4,8$ is the last quarter in the stratum of elapsed unemployment considered. Then, $\tau-(\tilde{u}-u)$ counts the quarters since start of treatment l yielding alignment of unemployment experience, because $u+\tau = \tilde{u}+(\tau-(\tilde{u}-u))$, and $Y^l(\tilde{u},\tau-(\tilde{u}-u))$ is the outcome of individuals who receive treatment l between period u and $\bar{u}$. For starts of l later than u, we have $\tilde{u}-u > 0$ and therefore, before l starts, $\tau-(\tilde{u}-u) < 0$. Then, these individuals are still unemployed, i.e. $Y^l(\tilde{u},\tau-(\tilde{u}-u)) = 0$ when the second argument of $Y^l(.,.)$ is negative. This way, we account for the fact that alternative treatments, for which the individual receiving treatment k in period u is eligible, might not start in the same quarter u.[15]

[15] Admittedly, the notation in equation (1.1) is cumbersome because we do not follow Sianesi (2003) and allow the alternative treatment l to start only in the same quarter u as treatment k. The problem is that one has to decide how to assign individuals who receive an alternative treatment later in the stratum considered. We think that program participation $l \in \{1,2,3\}$ in a later quarter $\tilde{u} > u$ (with $\tilde{u} < \bar{u}$) should not be interpreted as no participation (treatment 0) but rather we suggest to add such a case to the l–alternative for treatment k in quarter u. This is reflected in the definition of the parameter of interest.

The treatment parameter we actually estimate is the average within a stratum

$$\theta(k,l;\tau) = \sum_u g_u \theta(k,l;u,\tau) \ ,$$

with respect to the distribution g_u of starting dates u within the stratum.

Our estimated treatment parameter (1.1) mirrors the decision problem of the case worker and the unemployed who recurrently during the unemployment spell decide whether to start any of the programs now or to postpone participation to the future.

We evaluate the differential effects of multiple treatments assuming the following dynamic version of the conditional mean independence assumption (DCIA)

$$(1.2)\ E(Y^l(\tilde{u},\tau-(\tilde{u}-u))|T_u = k, u \leq \tilde{u} \leq \bar{u}, U \geq u-1, T_1 = ... = T_{u-1} = 0, X)$$

$$= E(Y^l(\tilde{u},\tau-(\tilde{u}-u))|T_{\tilde{u}} = l, u \leq \tilde{u} \leq \bar{u}, U \geq u-1, T_1 = ... = T_{u-1} = 0, X) \ ,$$

where X are time–invariant (during the unemployment spell) characteristics, $T_{\tilde{u}} = l$ indicates treatment l between u and $\bar{u}$ ($\bar{u}$ is the end of the stratum of elapsed unemployment considered), and $\tau \geq 0$, see equation (1.1) above and the analogous discussion in Sianesi (2004, p. 137). We effectively assume that conditional on X, conditional on being unemployed at least until period $u-1$, and conditional on not receiving any treatment before u (both referring to treatment in period u) individuals are comparable in their outcome for treatment l occurring between u and $\bar{u}$.

Building on Rosenbaum and Rubin's (1983) result on the balancing property of the propensity score in the case of a binary treatment, Lechner (2001) shows that the conditional probability of treatment k, given that the individual receives treatment k or treatment l, $P^{k|kl}(X)$, exhibits an analogous balancing property for the pairwise estimation of the ATT's of program k versus l. This allows to apply standard binary propensity score matching based on the sample of individuals participating in either program k or in program l. For this subsample, we simply estimate the probability of treatment k and then apply a bivariate extension of standard propensity matching techniques. Implicitly, we assume that the actual beginning of treatment within a stratum is random conditional on X.

To account for the dynamic treatment assignment, we estimate the probability of treatment k given that unemployment lasts long enough to make an individual 'eligible'. For treatment during quarters 1 to 2, we take the total sample of unemployed, who participate in k or l during quarters 1 to 2 (stratum 1), and estimate a Probit model for participation in k. This group includes those unemployed who either never participate in any program or who start some treatment after quarter 2. For treatment during strata 2 and 3, the basic sample consists of those unemployed who are still unemployed in the first month of the stratum.

We implement a stratified local linear matching approach by imposing that the matching partners for an individual receiving treatment k are still unemployed in the quarter before treatment k starts, i.e. we exactly align treated and nontreated individuals by elapsed unemployment duration in quarters. The expected counterfactual employment outcome for nonparticipation is obtained by means of a bivariate local linear regression on the propensity score and the starting month of the unemployment spell. We use a bivariate crossvalidation procedure to obtain the bandwidths in both dimensions (propensity score and beginning of unemployment spell) minimizing the squared prediction error for the average of the l–outcome for the nearest neighbors of the participants in program k.[16] An estimate for the variance of the estimated treatment effects is obtained through bootstrapping based on 200 resamples. This way, we take account of the sampling variability in the estimated propensity score.

As a balancing test, we use the regression test suggested in Smith and Todd (2005) to investigate whether the time–invariant (during the unemployment spell) covariates are balanced sufficiently by matching on the estimated propensity score $P^{k|kl}(X)$ using a flexible polynomial approximation. For each specification of the propensity score, the additional appendix reports the number of covariates for which the balancing test passes, i.e. the zero hypothesis is not rejected. Furthermore, we investigate whether treated and matched nontreated individuals differ significantly in their outcomes before the beginning of unemployment, in addition to those variables already used as arguments of the propensity score. We estimate these differences in the same way as the treatment effects after the beginning of the program. By construction, treated individuals and their matched counterparts exhibit the same unemployment duration until the beginning of treatment.

Finally, we need to discuss the plausibility of the DCIA (1.2) for our application.

[16] This method is an extension of the crossvalidation procedure suggested in Bergemann et al. (2004) and also used in Fitzenberger and Speckesser (2007). A detailed description of the implementation of the two dimensional bandwidth search can be found in Fitzenberger, Osikominu and Völter (2006).

As Sianesi (2004), we argue that the participation probability depends upon the variables determining re-employment prospects once unemployment began. Consequently, all individuals are considered who have left employment in the same two years (matching controls for beginning of unemployment) and who have experienced the same unemployment duration before program participation. Furthermore, observable individual characteristics and information from the previous employment spell have been included in the propensity score estimation. E.g., we consider skill information, regional information, occupational status, and industry which should be crucial for re-employment chances. Unfortunately, our data lack subjective assessments of labor market chances of the unemployed (e.g. by case workers). We argue that these are proxied sufficiently by the observed covariates in so far as they affect selection into the program. This is particularly plausible, since participation occurred at a fairly large scale, assignment was not very targeted and driven by the supply of programs, and case workers had little guidance on 'what works for whom'. Supporting our point of view, Schneider et al. (2006) argue that until 2002 assignment to training was strongly driven by the supply of available courses.[17]

1.4.2 Comparison to Alternative Dynamic Approaches

Abbring and van den Berg (2003), Lechner and Miquel (2005) as well as Lechner (2004), and Heckmann and Navarro (2007) propose three important alternative approaches to estimate treatment effects in a dynamic context. We now compare our approach building on Sianesi (2004), as described in the previous section, to these approaches. Under stronger assumptions than we are willing to make for our analysis, all three alternative approaches would allow to estimate more comprehensive treatment effects than estimated in this paper.

The timing–of–events approach by Abbring and van den Berg (2003) uses a continuous time duration model with unobserved heterogeneity, where time until treatment start and unemployment duration constitute two competing risks. The goal is to estimate the causal treatment effect on the hazard to leave unemployment. Identification of the causal effect of entering a program relies on the conditional randomness

[17]For the evaluation of the employment effects of job creation schemes in 1999/2000 based on administrative data for Germany, Caliendo et al. (2004) were able to use a survey asking about the motivation of participants (such information is not available for our data). It turned out that both using administrative data and controlling for these motivational variables did not result in noticeably different estimated program effects compared to using administrative data only. This evidence also supports our point of view.

of program starts and a non–anticipation condition as well as functional form assumptions involving e.g. a mixed proportional hazard model and a tight specification of the joint dependence between duration until treatment and the outcome variable unemployment duration. Our approach also considers the variation of starting dates during the unemployment spell, but relying on a selection on observables strategy, we estimate flexible discrete time hazards into the program where covariates are fully interacted with the elapsed duration of unemployment. By conditioning on elapsed unemployment duration by strata, we account for the endogenous selectivity of the group of individuals eligible for treatment at different points of time. In contrast to Abbring and van den Berg (2003), we allow for heterogeneity of treatment effects and our outcome variable employment is static. The two approaches have in common that the estimated treatment parameter corresponds to the effect of starting a program at a given point in time versus postponing it.

Our estimated treatment parameter (1.1) can be cast into the sequential treatment framework proposed by Lechner and Miquel (2005) and Lechner (2004). These studies consider the identification and estimation of dynamic treatment effects with matching methods in a context where selection into and out of programs takes place sequentially from one period or stage to the next. Lechner and Miquel distinguish two versions of the conditional independence assumption: strong and weak dynamic conditional independence. The strong version is inappropriate in our case, because it effectively rules out to match on elapsed unemployment duration which is affected by earlier treatments. Under the weak dynamic conditional independence assumption, it is possible to identify and estimate the effect of joining versus waiting for those who join at the period in question.[18] This assumption allows to identify other more comprehensive parameters as well. However, it involves common support requirements which are infeasible in our case because treatment or exit to employment at some point excludes later treatment.

Heckman and Navarro (2007) consider the semiparametric identification of dynamic treatment effects in structural dynamic discrete choice models. Similar to Abbring and van den Berg (2003), the treatment status is allowed to depend on unobserved factors in the outcome equation. Heckman and Navarro require the existence of instruments that affect choices but not outcomes for semiparametric identification

[18]For instance, the effect of training versus waiting in the second stratum for the group of participants corresponds to the following dynamic average treatment effect on the treated in the Lechner/Miquel framework: nonparticipation in period one and training in period two versus nonparticipation in both periods for the population of those who participate in the second period.

of causal effects. Matching methods, in contrast, rely on a rich set of conditioning variables that affect both the selection into treatment as well as the outcome such that any dependence between treatment assignment and outcome is netted out.[19] In a dynamic context, one needs instrumental variation at each stage of the sequential selection process or variation in the impact of time–invariant instruments (see Heckman and Navarro, 2007, Theorem 1). This variation must not be fully anticipated by the agents.[20] Using the reduced form model of Heckman and Navarro (section 2 of their paper), it is possible to identify the counterfactual outcome of waiting until a later period for those who join in an earlier period. Using the approach to identify other more comprehensive parameters in our case, more stringent modelling assumptions are necessary than we are willing to make and the necessary data requirements regarding instruments are not likely to hold. Unfortunately, we lack time–varying exogenous variables which affect the assignment process into treatment (see section 1.3 above).

1.5 Empirical Results

1.5.1 Estimation of Propensity Scores

Our empirical analysis is performed separately for the two samples of inflows from employment into unemployment, associated with transfer payment or program participation. To estimate the propensity scores, we run Probit regressions for training starting during the three time intervals for elapsed unemployment duration, i.e. 1–2 quarters (stratum 1), 3–4 quarters (stratum 2), and 5–8 quarters (stratum 3). Instead of estimating a multinomial choice model for entry in one of the three programs or no entry at all for each window of elapsed unemployment duration and sample, we estimate a series of binary Probit regressions. The additional appendix reports our preferred specifications for the 1986/87 and 1993/94 samples, which are obtained after extensive specification search.

The covariates considered are all defined for the beginning of unemployment and are

[19]Heckman, Ichimura, and Todd (1998) compare matching to conventional selection models and show how to exploit information on exclusion restrictions and additive separability of the outcome equation for matching. See Heckman and Navarro-Lozano (2004) for a comparison of matching, instrumental variables and control function methods in a static context.

[20]Abbring and van den Berg (2003, 2005) argue that it is often difficult to maintain exclusion restrictions in dynamic settings with forward-looking agents.

thus time–invariant for an individual during the unemployment spell. Personal characteristics considered are age (as dummies for five–year intervals), dummy variables for gender, marital status, having kids, being a foreigner and formal education (no vocational training degree, vocational training degree, tertiary education degree). In addition, we use information about the last employer, namely industrial sector and firm size dummies, and a number of characteristics of the previous job as employment status and information on earnings in the previous job. In particular, we use three variables containing information on earnings. Due to reporting errors and censoring, we do not know the exact earnings for all observations. Therefore, we distinguish the following three cases. First, we use a dummy variable that is equal to one if daily earnings are above 15 Euro (in 1995 Euros), roughly the minimum level to be subject to social security taxation.[21] Second, we have a dummy variable that indicates whether daily earnings are topcoded at the social security taxation threshold (*Beitragsbemessungsgrenze*). Third, we have a variable that records log daily earnings in the range between 15 Euro and the topcoding threshold and zero otherwise.

Regarding the employment and program participation history, we consider the following covariates. We use dummies indicating employment status in month 6, 12, and 24 before the beginning of unemployment. We also consider the number of months in regular employment during five years before the beginning of unemployment. The previous program participation history of an individual is captured by dummy variables that indicate participation in an ALMP program in year(s) 1, 2, and 3–5 before the beginning of unemployment. Differences in regional labor market conditions as well as supply of programs are the reason to include regional variables in the specification. We use the federal state of last employment and the unemployment rate as well as the population density at the district level. Finally, we also use the calendar month of the beginning of the unemployment period, either as a variable counting elapsed months since a given reference date (e.g. January 1960) or as dummies for the respective years and quarters.

Our specification search starts by using as many as possible of the covariates men-

[21] Monthly earnings below e.g. DEM 410 in 1986 and DEM 500 in 1992 in West Germany for marginal part–time employees (*geringfügig Beschäftigte*) were not subject to social security taxation and should therefore not be present in the data. In addition, it was possible to earn at most twice as much in at most two months of the year without contributing to the social insurance. Probably due to recording errors, the data shows a number of employment reports with zero or very low earnings. Since this information is not reliable, we only use the information for daily earnings reported above 15 Euro as a conservative cut–off point.

tioned above without interactions. The specification search is mainly led by the following two criteria: (i) single and joint significance, and (ii) balance of the covariates according to the Smith–Todd (2005) test. As regards the qualitative variables, like state, firm size and industry, which are split up into dummies for the different categories in the regression, we usually test for joint significance. When insignificance is found, the covariates are dropped. Furthermore, we test for the significance of interaction effects, in particular interactions with gender and age. In order to achieve balance of covariates, we test different functional forms (e.g. the square of a variable) and interaction effects. In a few cases, we keep insignificant covariates or interactions if they help to achieve balance. As we find the balancing test to be somewhat sensitive to small cell sizes we occasionally aggregate small groups that have similar coefficients. One example is the aggregation of two federal states.

The results for the Probit estimates show that the final specifications vary considerably over the inflow cohorts and the three time intervals even keeping the k/l–comparison constant.[22] On the one hand, this emphasizes the necessity to treat all 36 k/l–pairs separately. On the other hand, it makes it impossible to present and discuss all the specifications in detail. In general, the number of covariates decreases with elapsed unemployment duration. This is not surprising because many covariates contain information about the previous job, which should characterize someone in a better way who has only recently become unemployed compared to a long–term unemployed. Furthermore, since the 'better' types leave unemployment earlier, the long–term unemployed tend to be a more homogeneous subgroup.

Age effects are significant in most estimations. In particular, participants in retraining are younger than individuals in other groups. This reflects the assignment policy of the employment agency. The very comprehensive and expensive retraining schemes are preferably assigned to individuals who have a long time horizon of working life. Gender effects are also relevant in most cases, but they do not follow a common pattern. In cases where the foreign dummy is significant, it shows that foreigners have a lower probability to participate in any program. The employment history is important in most estimations. Previous participation in an ALMP program is sometimes significant. If so, it increases the probability of another program participation. The industrial sector of the previous job is sometimes significant and the firm size only rarely. In most estimations regional effects and the calendar date

[22] The tables with the propensity score estimations, the balancing tests and the figures showing the support of the propensity scores are displayed in the additional appendix.

of unemployment entry (seasonal effects) are contained.

For the balancing test in almost all cases, using a cubic in the estimated propensity score, we reject for at most one variable in the respective propensity score specification. Only in two out of 36 cases the test rejects for two variables. Considering both variants, i.e. the cubic and the quartic in the propensity score, the test does not reject for more than one variable in the specification in 20 out of 36 specifications. Overall, we are confident to have achieved a sufficient degree of balance between treatment and control groups in order for matching on the propensity score to be valid.

The graphical examination of the common support requirement for estimating the average treatment effect on the treated (ATT) for training versus waiting reveals that lack of common support is a problem only in some cases. In these cases, it occasionally happens, that for very small estimated participation probabilities there are only nonparticipants, but no participating counterparts with such low estimated participation probabilities. This poses no problem for estimating the treatment effect on the treated for treatment versus waiting. Only in two of the treatment versus waiting comparisons we excluded 1.3% and 1.8% of the treated observations, respectively, from the estimation. Overall, we are also satisfied with the overlap of support for treatment versus treatment (k/l–pairs with $k, l \geq 1$). Though the graphical inspection of common support seems to reveal slight differences in support in a few cases, these differences mostly lie within the close neighborhood of the respective treated observation determined by the bandwidth. Therefore, we proceed without restricting the samples except for four cases where we drop about 2% to 4% of the treated observations, respectively. Detailed results of the balancing tests and common support graphics are shown in the additional appendix.

1.5.2 Estimated Treatment Effects

The outcome variable is the average of monthly employment dummies in a quarter. We match participants in treatment k and participants in treatment l by their similarity in the estimated propensity scores[23] and the starting month of the unemployment spell. For matching, we use only eligible participants in l who are still unemployed in the quarter before treatment starts and we align them by the quar-

[23] We use the fitted index $X_i\hat{\beta}$ from the Probit estimates.

ter of elapsed unemployment duration. The ATT is then estimated separately for quarters $\tau = 0, ..., \tau_{\max}$ since the beginning of program k, where $\tau_{\max} = 31, 29, 25$, respectively, for stratum 1, 2, and 3. The expected counterfactual employment outcome for l is obtained by means of a local linear regression on the propensity score and the starting month of the unemployment spell among the eligible l–group. We obtain an estimate for the variance of the estimated treatment effects through bootstrapping the entire observation vector for a spell in our inflow sample. This way, we take account of possible autocorrelation in the outcome variable. Inference is based on 200 resamples. As a further test of the matching quality, we estimate in the same way the differences between participants and matched nonparticipants during quarters 1 to 8 before the beginning of unemployment. By construction, participants in k and matched eligible members of the l–group are unemployed between the beginning of their unemployment spell and the beginning of the treatment in the k–group.

Figures 1.2–1.7 graphically represent the evaluation results. Each figure contains a panel of three times three graphs, where each row represents one pairwise comparison of two treatments and each column represents one of the three intervals of elapsed duration of unemployment at the beginning of the treatment, i.e. 1–2 (stratum 1), 3–4 (stratum 2), or 5–8 (stratum 3) quarters since the start of the unemployment spell. The graphs display the estimated average treatment effect for the treated during quarters 0 to $\tau_{\max}$ since the beginning of the treatment and the differences during 8 quarters before the beginning of the unemployment spell. We put pointwise 95%–confidence intervals around the estimated treatment effects. The vertical gap at zero reflects the variable length of time between the start of the unemployment spell and the start of the treatment.

In order to contrast the initial negative lock–in effects of the programs with the later positive program effects, we calculate the cumulated effects of every program 8, 16, and 24 quarters after the beginning of the program. The cumulated effects ($\equiv$ sum of quarter specific treatment effects) are calculated as the sum of the effects depicted in figures 1.2–1.7 starting in quarter 0 and summing up over the first 8, 16, and 24 quarters, respectively. Tables 1.7 and 1.9 provide the results. These effects show the change in the total number of quarters in employment since the beginning of treatment. When the cumulated effects become positive then a negative lock–in effect is compensated by positive effects afterwards. The estimated standard errors are based on the bootstrap covariance estimates for the quarter specific treatment effects. A potential drawback of considering cumulated effects is that many of them are rather imprecisely estimated because they are summed over all quarters such that negative short–run effects and positive medium– to long–run effects are lumped together. Therefore, we also include a table (table 1.8) for training versus waiting with average ATT's, that are averages of the quarter specific treatment effects during the first three years and from year four onwards after the beginning of treatment. Table 1.8 allows to assess in a parsimonious way whether persistent significantly positive effects exist after the end of the lock–in period.

Training versus Waiting

We first discuss the effects of the three training programs against the alternative of waiting, i.e. no treatment during the time interval (stratum) of elapsed unemployment duration, displayed in figures 1.2 (cohort 86/87) and 1.5 (cohort 93/94).

We do not find significant pre–unemployment employment differences in any case. Since all individuals become unemployed eventually, this test for matching quality should focus on the differences during the earlier quarters. There is no evidence of systematic differences in employment rates between treated and associated matched individuals. This suggests that time–invariant unobserved heterogeneity does not invalidate our matching approach.

The results for 86/87 in figure 1.2, show positive medium–run (1–3 years) and long–run (4–6 years) post treatment effects of all three training programs after a negative lock–in effect in the program right after the beginning of treatment. These effects are typically of the magnitude 10 to 20 percentage points (ppoints) and prove significant. They are smaller and not significant for PF in the second and third stratum. For SPST and RT the medium–run effects lie even above 20 ppoints for strata 2 and 3 and are larger than the long–run effects. As expected, the lock–in periods are shortest for PF (typically the shortest treatment), lasting at most 3 quarters, and longest for RT, lasting up to two years. SPST lies in between for strata 1 and 2 with a lock–in period of about 1 year and shows a very short lock–in period of 2 quarters for stratum 3. The positive effects for SPST show similar patterns for the three strata (similar to the results for SPST in Fitzenberger and Speckesser, 2007), with the effects being slightly higher in strata 2 and 3. For RT the positive medium–run effects are larger for strata 2 and 3 compared to stratum 1 and the long–run effects are larger for stratum 2 compared to both strata 1 and 3.

For the 93/94 cohort, figure 1.5 shows similar patterns for training versus waiting. For PF, we find shorter lock–in periods for strata 2 and 3 and small positive but insignificant treatment effects after the lock–in period in stratum 1. For strata 2 and 3, we now find significantly positive medium– and long–run treatment effects of 10 to 15 ppoints. Again, the lock–in period is longer for SPST and even longer for RT. The significantly positive medium– and long–run effects for SPST lie between 10 and 20 ppoints and are more persistent than for the earlier cohort. The positive medium– and long–run effects for RT in stratum 1 are below 10 ppoints and barely significant. The effects are somewhat stronger for strata 2 and 3.

Next, we discuss the cumulated effects of the different programs against the alternative of waiting, which are reported in table 1.7. This allows for a simple comparison of the ATT effects across programs, though it is important to recall that these effects for the treated cannot be compared because they are based on the separate groups of participants in the different programs. It will be interesting to contrast these effects to the results of the pairwise program comparisons reported in the next subsection.

For the 86/87 cohort, the cumulated long–run effects after 24 quarters are significantly positive at the 10%–level for all cases, except PF in stratum 3. Overall, SPST shows the largest long–run effects with the highest value of 4.2 in stratum 3, i.e. during the 24 quarters after the beginning of the treatment the treated individuals are employed on average for about 4 quarters more than had they not been treated. For SPST and RT, the long–run effects are higher in later strata, though one can not put a causal interpretation to this because the selection of individuals in the different strata changes. There are less cases with significantly positive cumulated effects after 16 quarters. After 8 quarters, the cumulated effects are still negative for RT due to the longer lock–in period, mostly positive for SPST and PF, and significantly positive in strata 2 and 3 for SPST.

For the 93/94 cohort, the cumulated long–run effects after 24 quarters are significantly positive in all strata for SPST, in strata 2 and 3 for PF, and in stratum 2 for RT. For SPST, the pattern is similar to the earlier cohort. For PF, the effect is higher in strata 2 and 3 and much lower in stratum 1. Also for RT, the effects are lower and even significantly negative in stratum 1. Early treatments for PF or RT in stratum 1 show worse effects for 93/94 compared to 86/87. The effects at 8 and 16 quarters for RT show stronger lock–in effects for the later cohort. For PF in strata 2 and 3, there are stronger positive effects already at 8 and 16 quarters.

Table 1.8 shows the yearly averages of the ATT's which are typically more precisely estimated than the quarter specific treatment effects and the cumulated effects. For SPST, all average ATT's are significantly positive from year 2 onwards and slightly smaller for the cohort 93/94. For RT, all effects from year 3 onwards are significantly positive for the cohort 86/87 and only from year 4 onwards significantly positive for the cohort 93/94. RT shows longer significantly negative lock–in effects for the cohort 93/94 compared to 86/87. The pattern of the estimated effects for PF is less clear cut. For the cohort 86/87, there are significantly positive effects for one or two of the periods following year 1. For 93/94, PF shows significantly positive effects for all periods following year 1 in strata 2 and 3 but no significantly positive effects in stratum 1.

Summing up, our results on training versus waiting show that most training programs yield significantly positive and fairly persistent medium– and long–run treatment effects. There are strong lock–in effects, with RT showing the longest lock–in periods (up to 8 quarters). The cumulated effects and the average treatment effects during years 2 to 4 are significantly positive for most programs. Overall, SPST seems to show the best results for the treated individuals. The positive effects of SPST deteriorate little for the later cohort. For RT, there is a noticeable increase in the lock–in period and a noticeable decline in the treatment effects for the later cohort and, for PF, the treatment effects deteriorate for stratum 1 and improve for later program starts.[24] The slight deterioration of the treatment effects for the later cohort could be caused by the worse business cycle conditions in the 90s.

Pairwise Comparisons of Training Programs

Next, we estimate pairwise ATT's both of the treatment k versus the alternative l for the treated in k and the treatment l versus k for the treated in l. As mentioned above, the first ATT does not necessarily coincide with the negative of the second ATT because of effect heterogeneity and the different composition of the two treatment groups (Lechner, 2001). The pairwise comparison allows to investigate whether the different programs are well targeted on average. With individual heterogeneity of

[24]Following the suggestion of a referee, we also investigated whether there are heterogeneous treatment effects by gender, age, and qualification. In the matched samples, we regressed outcome differences on these covariates. However, based on bootstrapped standard errors we did not find any significant differences. These results are available upon request.

treatment effects, it could very well be the case that the participants in SPST fare better on average through participating in SPST as compared to RT even though the participants in RT also fare better on average through participating in RT as compared to SPST. This example is used because we find some evidence for such effects, though they often are not significant.

The quarterly treatment effects for the pairwise comparisons are displayed in figures 1.3, 1.4, 1.6, and 1.7. After a short description of these effects, our discussion focuses on the cumulated effects in table 1.9. Note that for the pairwise comparisons, the control groups used for local linear matching are considerably smaller compared to evaluating one training program versus nonparticipation, see tables 1.3 and 1.4.

In the vast majority of cases, we do not find significant pre–unemployment employment differences. In a small number of cases, there are significant (but barely so) employment differences for some quarters before the beginning of unemployment.[25] Therefore, we conclude that there are no systematic differences in employment rates left between treated and associated matched individuals.

We find significant short–run treatment effects in a number of cases reflecting the different lock–in periods of the three training programs. RT performs worse than the two other programs during the first two years and PF tends to perform better during the first year. However, we do not find this for all cases. We do not find persistent medium– and long–run effects. In a number of cases, the treatment effects in the medium and long run are significant over a short time period and display quite erratic movements.

The estimated cumulated effects in table 1.9 suggest that for the cohort 86/87 most significant effects are caused by the differential lock–in periods. Comparing SPST with RT for those treated in SPST ('SPST vs RT'), we find strong significantly positive effects after 8 quarters in stratum 1 and 3. Comparing RT with SPST and RT with PF both for those treated in RT, we find no significantly positive effects and the point estimates are even negative in a number of cases. For participants in SPST, SPST seems to outperform RT at 16 quarters for strata 1 and 3, but the cumulated effects are reduced at 24 quarters and not significant any more. For participants in RT, SPST seems to outperform RT as well at 16 quarters for stratum 1 but again the effect at 24 quarters is reduced and not significant any more. PF seems to

[25]These differences in employment history often become insignificant, if larger bandwidths are used. Further details are available upon request.

outperform SPST for participants in SPST in stratum 1 after 24 quarters, whereas the cumulated long–run effects are insignificant for participants in PF. The long–run cumulated effects for RT in comparison to PF for participants in RT are positive and sizeable in stratum 2 and 3, but not significant. The long–run cumulated effects of PF in comparison to RT are also positive in stratum 1 and 3 but not significant.

For the cohort 93/94, the cumulated effects at 8 quarters are qualitatively similar reflecting again the different lock–in periods. Both PF, for stratum 1 and 3, and SPST, for all strata, seem to outperform RT in the short and medium run for the participants in PF and SPST, respectively. In the long run we only find significant effects for participation in SPST compared to RT in stratum 1. RT is also outperformed by SPST and PF even for participants in RT, though the effects are only strongly significant at 8 and at 16 quarters (the effects are of similar size at 24 quarters). Comparing SPST and PF, the cumulated effects are not significant but the point estimates suggest that SPST outperforms PF at least for the own participants.

Summing up, our results on the pairwise comparisons are much weaker compared to the comparison of training versus waiting, because the standard errors for the pairwise comparisons are much higher. Nevertheless, we can draw some conclusions. The significant cumulated effects after 8 quarters reflect the different lock–in periods for the three training programs. Most medium– and long–run cumulated effects are insignificant which suggests that in these cases, the employment outcome of the treated individuals could not have been improved on average in the medium or long run by reallocating them to a different training program. There is, however, some evidence for SPST and PF outperforming RT in the medium and long run even for the participants in RT, for the 93/94 cohort. The point estimates for SPST versus PF suggest for stratum 1 in 86/87 that the cumulated employment effect would have been better, if participants in SPST had instead participated in PF. For 93/94, the point estimates suggest that SPST outperforms PF in the medium and the long run even for participants in PF. However, none of these effects for 93/94 are significant.

1.6 Conclusions

Based on a unique administrative data set, which has only recently become available, we analyze the long–run employment effects of three types of public sector sponsored training in West Germany, which do not involve a job for the participants. The three types of training are Practice Firm (PF), Retraining (RT), and the Provision of Specific Professional Skills and Techniques (SPST). Specifically, we estimate the average treatment effect on the treated (ATT) against the alternative of nonparticipation in any program as well as for pairwise comparisons among the three programs. We take inflow samples into unemployment for West Germany in 1986/87 and 1993/94. We use the approach for multiple treatment evaluation suggested by Lechner (2001) and Imbens (2000) and apply it to a dynamic setup. Slightly modifying the approach suggested by Sianesi (2003, 2004), we distinguish three types of treatment depending upon the elapsed duration of unemployment when treatment starts, i.e. treatment starts during the first two quarters (stratum 1), during the third or fourth quarter (stratum 2), and between the fifth and the eighth quarter (stratum 3).

When comparing treatment against nonparticipation, the estimated treatment effects in almost all cases involve first a lock–in period with negative treatment effects and significantly positive treatment effects in the medium and long run. The lock–in period is shortest for PF (at most 2 quarters) and longest for RT (around 2 years). SPST lies in between with a lock–in period of around 4 to 6 quarters. The treatment effects deteriorate slightly from 1986/87 to 1993/94 in a number of cases, especially for RT and especially for treatments starting in stratum 1. For RT, the length of the lock-in period increases considerably for the later cohort. Both could reflect the worse business cycle conditions in the 1990s. The cumulated effects are significantly positive for most programs.

The pairwise comparisons of the three treatments, one against another, show first the differences in the lock–in periods and in most cases insignificant treatment effects in the medium and long run. There is, however, some evidence for SPST and PF outperforming RT in the medium and long run for the 1993/94 cohort. For 1993/94, SPST tends to outperform PF, but the effect is not significant.

Overall, SPST shows the best results for the treated individuals and the positive treatment effects for SPST are almost at the same level for 1993/94 compared to 1986/87. Note that SPST is by far the largest program and its share is even higher in 1993/94 compared to 1986/87. It is remarkable how little the effectiveness of SPST differs between the two time periods despite the differences in business cycle conditions and the apparent change in the timing and length of treatments.

In comparison to the study by Lechner et al. (2005a) based on the same data source, our general results for the 1993/94 cohort are quite similar in most cases, even though the exact treatment definition, the choice of valid observations, and the employed econometric methods differ substantially. Notable differences from the results reported in Lechner et al. (2005a) are that we find significantly positive effects for treatments relative to nonparticipation much earlier after the treatment starts and that our results for RT in comparison to other training programs are often negative.

Our study draws a somewhat more positive picture of large scale public sector sponsored training programs compared to the previous literature. However, an overall assessment of the microeconomic effects is not possible since various necessary information for a comprehensive cost–benefit–analysis are lacking in our data. Since the relative performance of SPST tends to improve over time and PF does not seem to dominate the other two programs, our evidence is in contrast to some of the conclusions in the surveys by Martin and Grubb (2001), Kluve and Schmidt (2002), and OECD (2005) advocating a strong on–the–job component for public sector sponsored training to show positive employment effects.

1.7 References

Abadie, A. and G. Imbens (2006). "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74, 235-267.

Abbring, J. and G.J. van den Berg (2003). "The Nonparametric Identification of Treatment Effects in Duration Models." *Econometrica* 71, 1491–1517.

Abbring, J. and G. van den Berg (2005). "Social Experiments and Instrumental Variables with Duration Outcomes." Tinbergen Institute Discussion Paper No. 05-047/3.

Bender, S., A. Bergemann, B. Fitzenberger, M. Lechner, R. Miquel, S. Speckesser, C. Wunsch (2005). "Über die Wirksamkeit von Fortbildungs- und Umschulungsmaßnahmen." *Beiträge zur Arbeitsmarkt- und Berufsforschung* 289, IAB, Nürnberg.

Bender, S., A. Haas, and C. Klose (2000). "IAB employment subsample 1975–1995." *Schmollers Jahrbuch (Journal of Applied Social Science Studies)* 120, 649–662.

Bender, S. and C. Klose (2000). "Berufliche Weiterbildung für Arbeitslose – ein Weg zurück in die Beschäftigung? Analyse einer Abgängerkohorte des Jahres 1986 aus Maßnahmen der Fortbildung und Umschulung mit der ergänzten IAB–Beschäftigtenstichprobe 1975-1990." *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung* 33, 421–444.

Bergemann, A. B. Fitzenberger, and S. Speckesser (2004). "Evaluating the Dynamic Employment Effects of Training Programs in East Germany Using Conditional Difference–in–Differences." ZEW Discussion Paper No. 04-41.

Bundesanstalt für Arbeit (1987, 1992, 1997). *Förderung der beruflichen Bildung 86, 91, Berufliche Weiterbildung 96.* Nürnberg: Bundesanstalt für Arbeit (various issues).

Caliendo, M., R. Hujer, S.L. Thomsen, and C. Zeiss (2004). "Einfluss von motivationalen Merkmalen aus Befragungsdaten auf die Schätzung des Beschäftigungseffektes von ABM–Maßnahmen." Unpublished Manuscript, Goethe University Frankfurt.

Fay, R. (1996). "Enhancing the Effectiveness of Active Labour Market Policies: Evidence from Programme Evaluations in OECD countries." Labour Market and Social Policy Occasional Papers, 18, OECD, Paris.

Fitzenberger, B. and H. Prey (2000). "Evaluating Public Sector Sponsored Training in East Germany." *Oxford Economic Papers* 52, 497–520.

Fitzenberger, B., A. Osikominu and R. Völter (2006). "Get Training or Wait? Long–Run Employment Effects of Training Programs for the Unemployed in West Germany." ZEW Discussion Paper No. 06-039.

Fitzenberger, B. and S. Speckesser (2007). "Employment Effects of the Provision of Specific Professional Skills and Techniques in Germany." *Empirical Economics* 32(2), 529-573.

Fredriksson, P. and P. Johansson (2003). "Program Evaluation and Random Program Starts." Institute for Labour Market Policy Evaluation (IFAU), Uppsala, Working Paper, 2003:1.

Fredriksson, P. and P. Johansson (2004). "Dynamic Treatment Assignment – The Consequences for Evaluations Using Observational Data." IZA Discussion Paper No. 1062.

Hagen, T. and V. Steiner (2000). "Von der Finanzierung der Arbeitslosigkeit zur Förderung der Arbeit – Analysen und Handlungsempfehlungen zur Arbeitsmarktpolitik." ZEW Wirtschaftsanalyszen, 51, Nomos, Baden-Baden.

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

Heckman, J. H. Ichimura, J.A. Smith and P. Todd (1998). "Characterizing Selection Bias using Experimental Data." *Econometrica* 65, 1017–1098.

Heckman, 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*, Vol. 3 A, Amsterdam: Elsevier Science, 1865–2097.

Heckman, J. and S. Navarro-Lozano (2004). "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic choice Models." *Review of Economics and Statistics* 86, 30–57.

Heckman, J. and S. Navarro (2007). "Dynamic Discrete Choice and Dynamic Treatment Effects." *Journal of Econometrics* 136, 341–396.

Ichimura, H. and O. Linton (2001). "Asymptotic Expansions for some Semiparametric Program Evaluation Estimators." Discussion paper, London School of Economics and University College London.

Imbens, G. (2000). "The Role of the Propensity Score in Estimating Dose–Response Functions" *Biometrika* 87, 706-710.

Kluve, J. and C. Schmidt (2002). "Can training and employment subsidies combat European unemployment?" *Economic Policy* 35, 441–448.

Lechner, M. (2000). "An evaluation of public sector sponsored continuous vocational training programs in East Germany." *Journal of Human Resources* 35, 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.) (2000), *Econometric Evaluation of Active Labor Market Politics in Europe*, Heidelberg: Physica–Verlag.

Lechner, M. (2004). "Sequential Matching Estimation of Dynamic Causal Models." Discussion Paper 2004–06, University of St. Gallen.

Lechner, M. and R. Miquel (2005). "Identification of the Effects of Dynamic Treatments by Sequential Conditional Independence Assumptions" Discussion Paper, University of St. Gallen.

Lechner, M., R. Miquel, and C. Wunsch (2005a). "Long–Run Effects of Public Sector Sponsored Training in West Germany." IZA Discussion Paper No. 1443.

Lechner, M., R. Miquel, and C. Wunsch (2005b). "The Curse and Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany after Unification." Discussion Paper, University of St. Gallen.

Martin, J.P. and Grubb, D. (2001). "What works and for whom: A review of OECD countries' experiences with active labour market policies." *Swedish Economic Policy Review* 8, 9–56.

OECD (2005). "Labour Market Programmes and Activation Strategies: Evaluating the Impacts." Chapter 4 of Employment Outlook, OECD, Paris.

Plassmann, G. (2002). "Der Einfluss der Arbeitslosenversicherung auf die Arbeitslosigkeit in Deutschland — eine mikroökonomische und empirische Untersuchung." *Beiträge zur Arbeitsmarkt- und Berufsforschung* 255, IAB, Nürnberg.

Rosenbaum, P.R. and D.B. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70, 41–55.

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

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

Schneider, H., K. Brenke, D. Hess, L. Kaiser, J. Steinwede und A. Uhlendorff (2006). "Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission – Modul 1b: Förderung beruflicher Weiterbildung und Transferleistungen." IZA Research Report, No. 7, Bonn.

Sianesi, B. (2003). "Differential Effects of Swedish Active Labour Market Programs for Unemployed Adults in the 1990s." Discussion Paper, Institute for Fiscal Studies, London.

Sianesi, B. (2004). "An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s." *Review of Economics and Statistics* 86, 133–155.

Smith, J.A. and P. Todd (2005). "Rejoinder." *Journal of Econometrics* 125, 365–375.

Speckesser, S. (2004). "Essays on Evaluation of Active Labour Market Policy." Dissertation, University of Mannheim.

Steffen, J. (2005). "Sozialpolitische Chronik Arbeitslosenversicherung seit 1969." Arbeitnehmerkammer Bremen, Bremen.

Wunsch, C. (2006). "Labour Market Policy in Germany: Institutions, Instruments and Reforms since Unification." Discussion Paper, University of St. Gallen.

1.8 Appendix

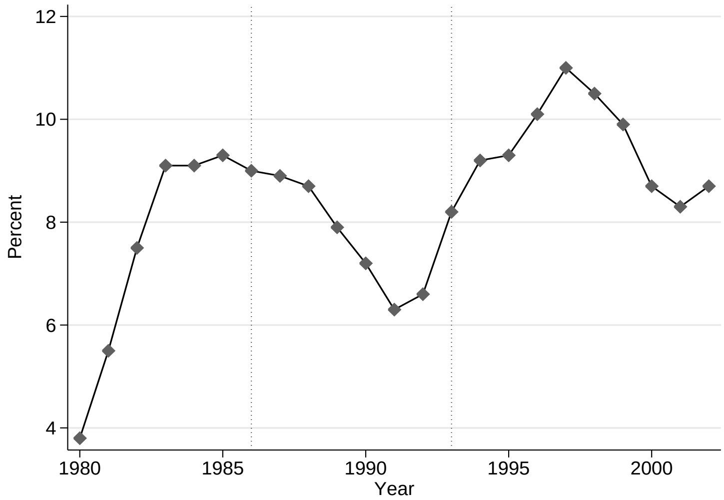
1.8.1 Descriptive Statistics and Description of Data

Table 1.2: Participation in Further Training in West Germany until 1997

Year	Annual entries				Annual average stocks
	Total	Further vocational training	Retraining	Integration subsidy	
1980	247.0	176.5	37.9	32.6	177.1
1985	409.3	336.5	45.1	27.7	245.8
1986	530.0	426.0	59.1	44.9	308.1
1987	596.3	482.6	64.5	49.2	346.1
1988	565.6	448.7	65.7	51.2	361.5
1989	489.9	388.4	61.0	40.8	357.9
1990	574.0	442.8	63.4	67.9	349.7
1991	593.9	474.5	70.5	48.9	364.5
1992	574.7	464.5	81.5	28.7	372.1
1993	348.1	266.0	72.2	9.9	348.4
1994	306.8	224.9	73.1	8.8	308.8
1995	401.6	309.7	81.8	10.0	304.3
1996	378.4	291.6	77.3	9.5	306.6

Remark: All numbers in thousands. Source: Bundesanstalt für Arbeit (1987, 1992, 1997).

Figure 1.1: Unemployment Rate in West Germany



Source: Federal Employment Agency

Table 1.3: Participation in First Training Program for the Inflow Samples into Unemployment

Training Program	Frequency	Percent of inflow sample	Percent among treated
Cohort 86/87			
Practice Firm	246	1.2	14.4
SPST	1,093	5.2	63.8
Retraining	375	1.8	21.9
No training program above	19,188	91.8	–
Total inflow sample	20,902	100	100
Cohort 93/94			
Practice Firm	325	1.3	11.9
SPST	1,944	7.8	71.3
Retraining	458	1.8	16.8
No training program above	22,324	89.1	–
Total inflow sample	25,051	100	100

Remark: Programs that start before a new job is found are considered. We exclude training programs which start together with a job (like integration subsidies) or which involve a very small number of participants since they are not targeted on inflows into unemployment (as career advancement and German language courses). Furthermore, we do not consider the very short programs according to §41a of the Labor Promotion Act, which are only offered to the 1986/87 inflow sample as separate programs, but treat them as open unemployment. This improves the comparability of the inflow samples since comparable very short–term programs offered to the 1993/94 inflow sample are not recorded as programs but as open unemployment in our data. Thus, a participation in retraining after a §41a program is counted as the first program.

Table 1.4: Number of Training Spells and Length of Unemployment before Program Start

	Practice Firm		SPST		Retraining	
	86/87	93/94	86/87	93/94	86/87	93/94
1–2 quarters	74	102	503	528	172	198
3–4 quarters	60	102	257	481	101	138
5–8 quarters	69	86	176	669	71	106
>8 quarters	43	35	157	266	31	16
Total	246	325	1,093	1,944	375	458

Remark: The time intervals indicate the quarter of program start relative to the beginning of the unemployment spell.

Table 1.5: Elapsed Duration of Unemployment in Months at Beginning of Training Spell

	Practice Firm		SPST		Retraining	
	86/87	93/94	86/87	93/94	86/87	93/94
Average	15.8	11.4	13.3	12.9	10.2	8.1
25%–Quantile	5	5	3	5	3	3
Median	10	9	6	11	6	7
75%–Quantile	19	15	14	18	12	12

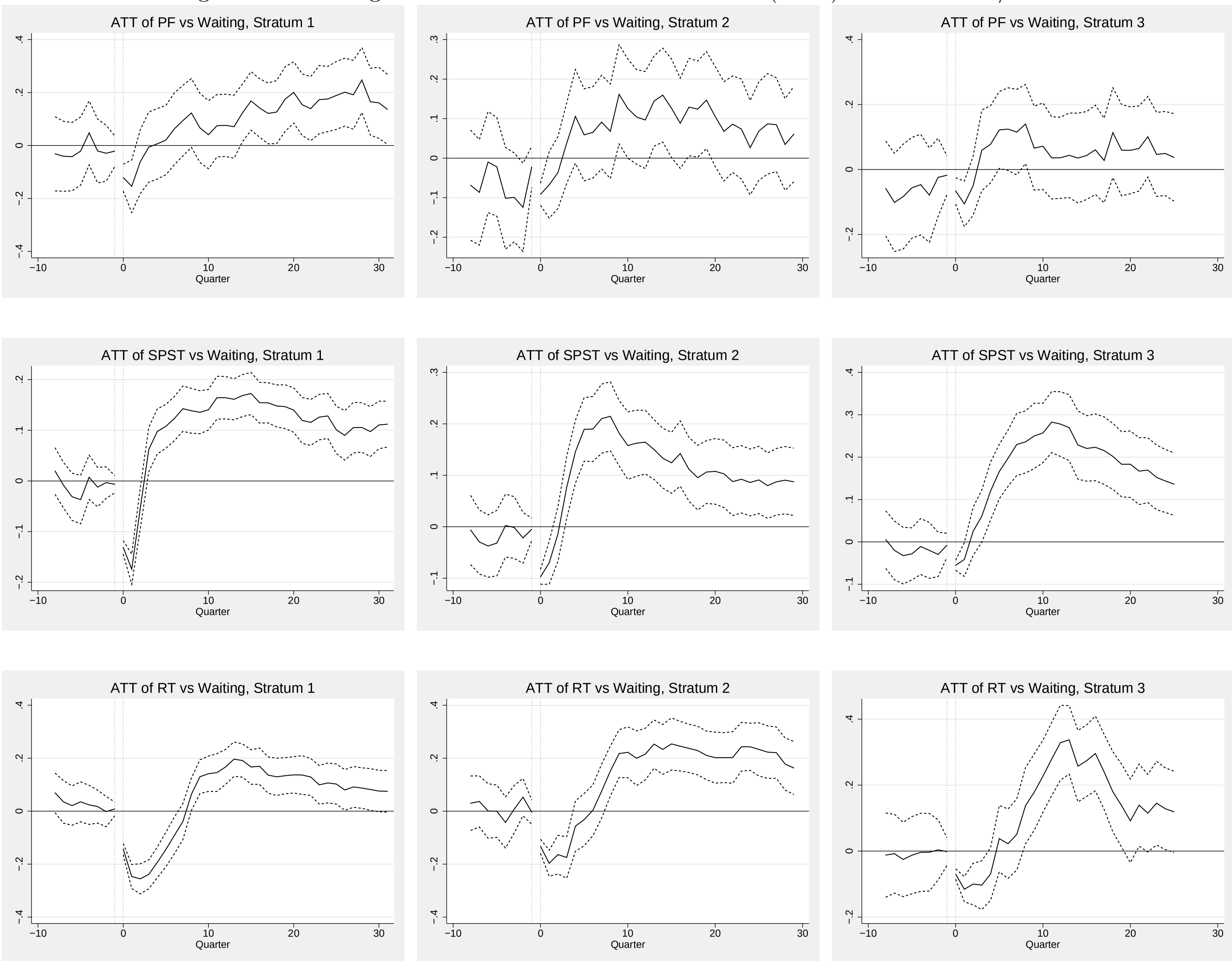
Table 1.6: Realized Duration of Training Spells in months

	Practice Firm		SPST		Retraining	
	86/87	93/94	86/87	93/94	86/87	93/94
Average	5.1	5.7	4.9	6.3	13.1	14.9
25%–Quantile	2	3	2	3	5	6
Median	5	6	4	6	12	16
75%–Quantile	6	8	7	8	22	21

Remark: The duration of the training spell is defined as the number of months of uninterrupted training.

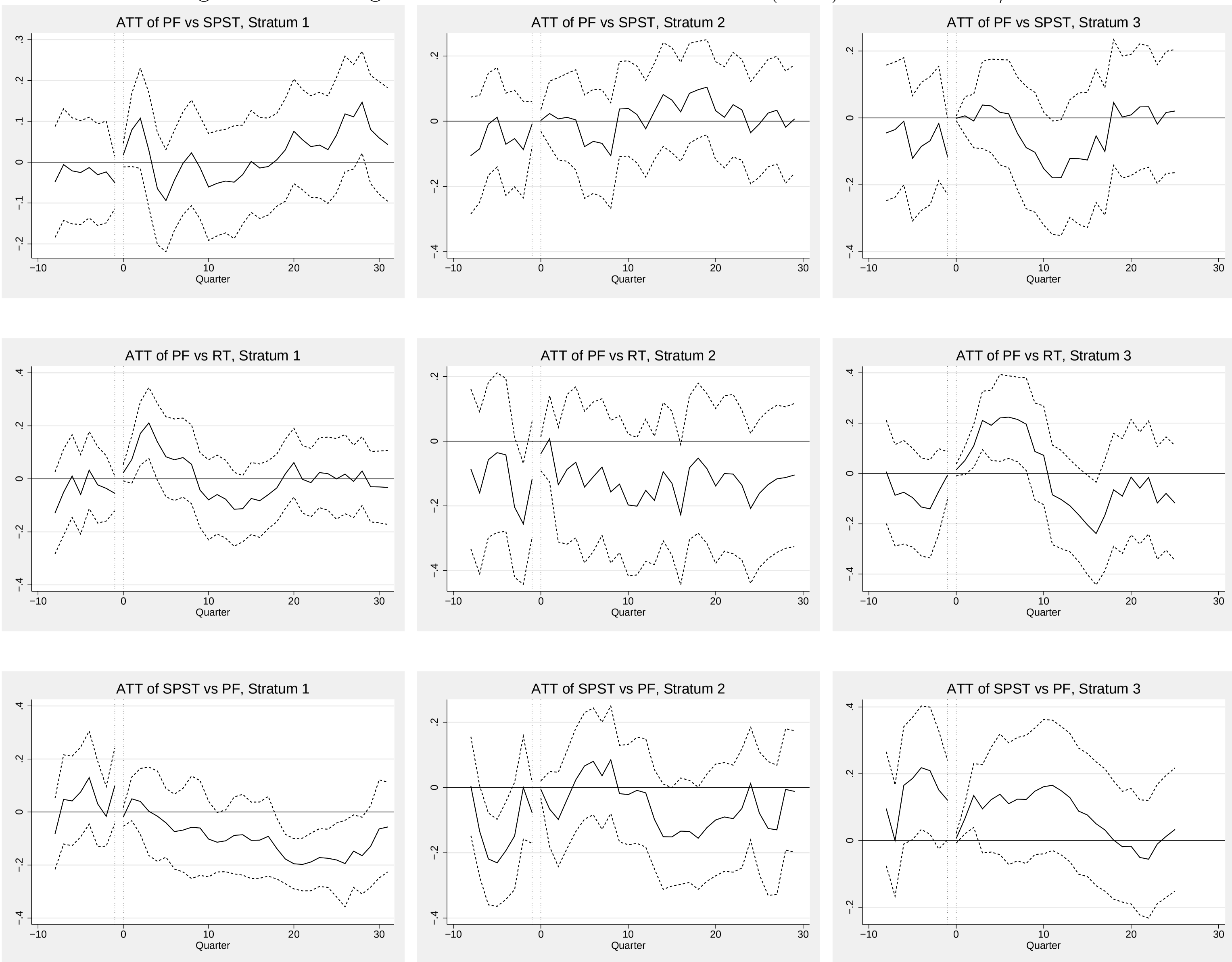
1.8.2 Estimated Employment Effects of Further Training Programs

Figure 1.2: Average Treatment Effect on the Treated (ATT) for Cohort 86/87



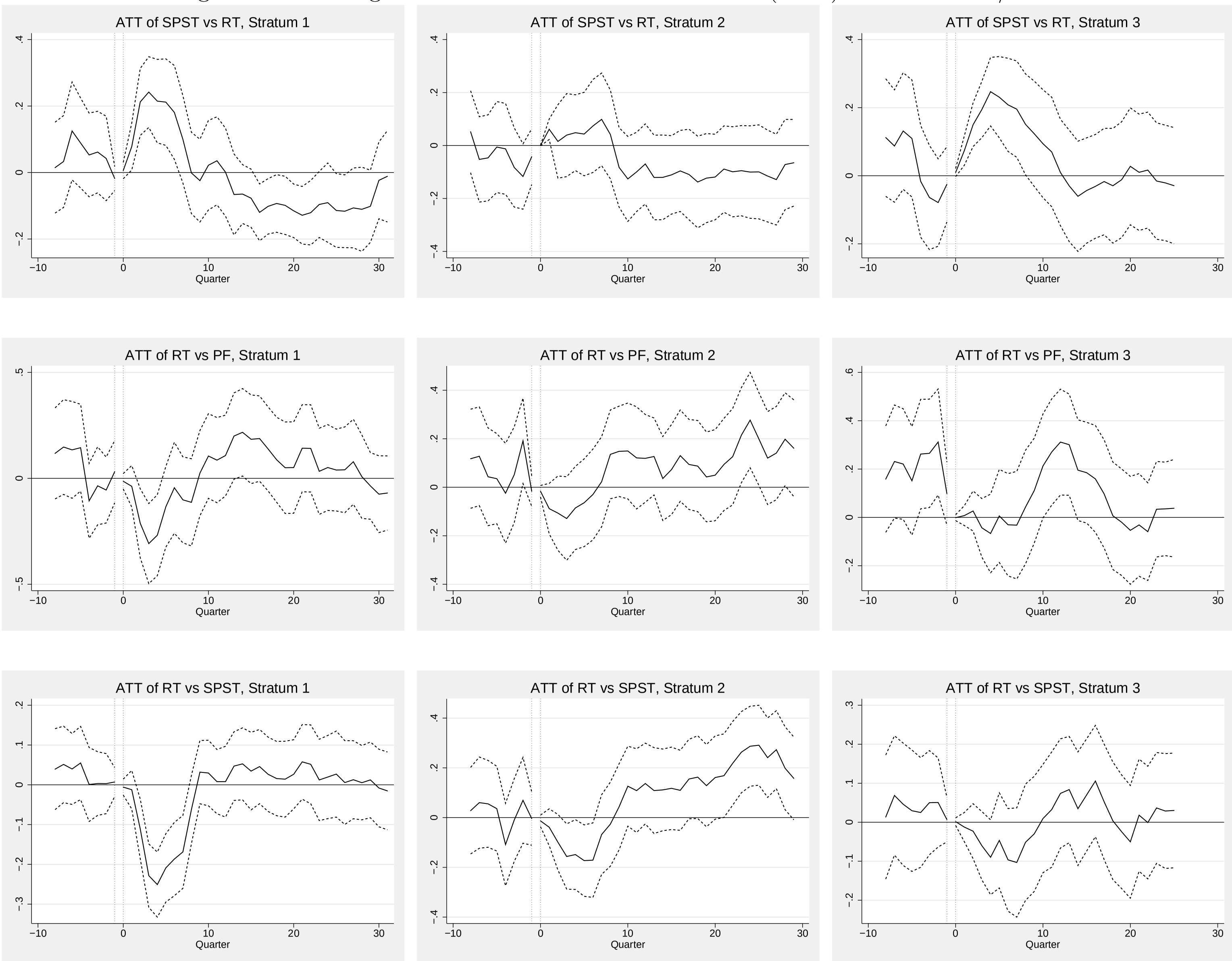
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) quarters on the abscissa.

Figure 1.3: Average Treatment Effect on the Treated (ATT) for Cohort 86/87



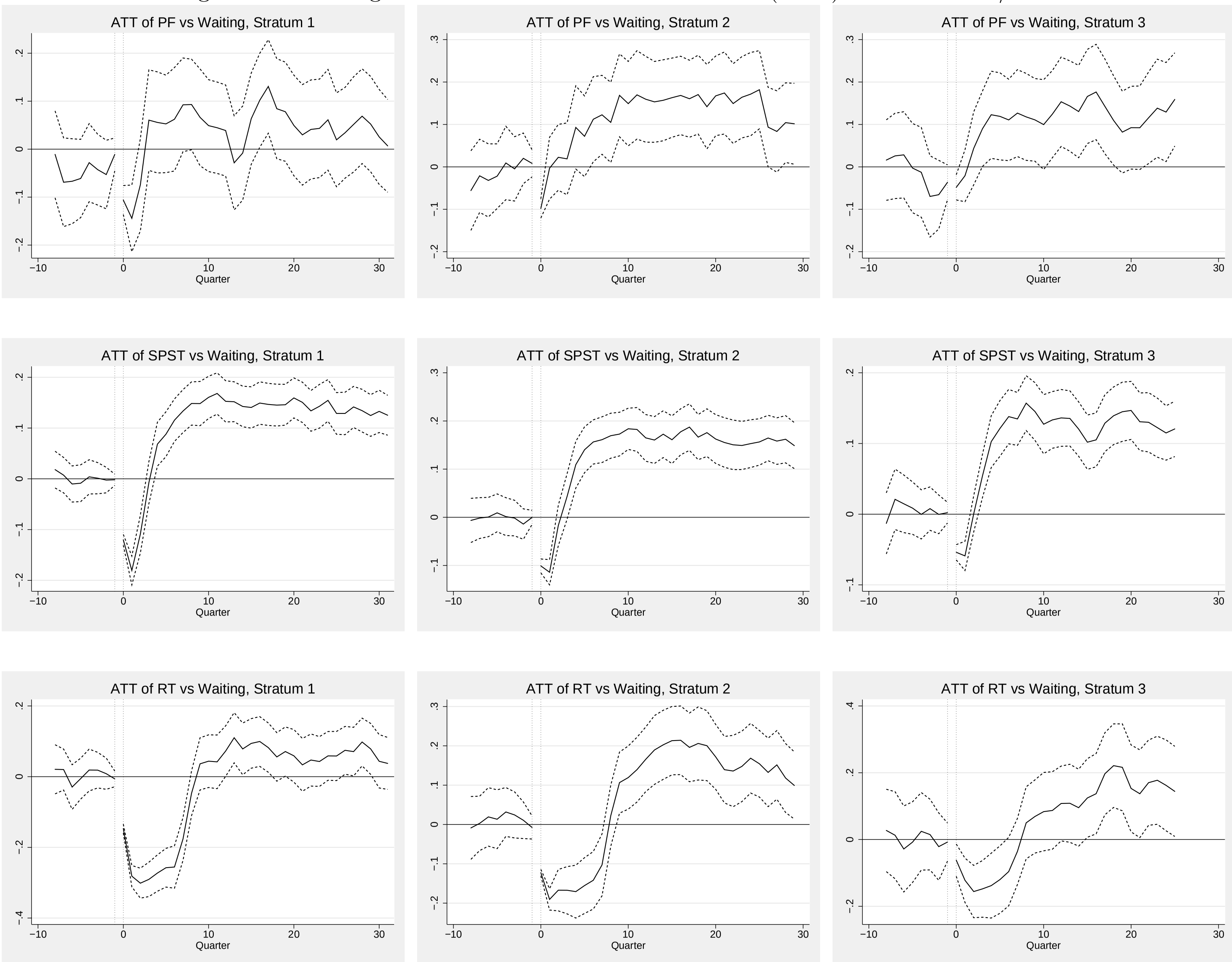
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) quarters on the abscissa.

Figure 1.4: Average Treatment Effect on the Treated (ATT) for Cohort 86/87



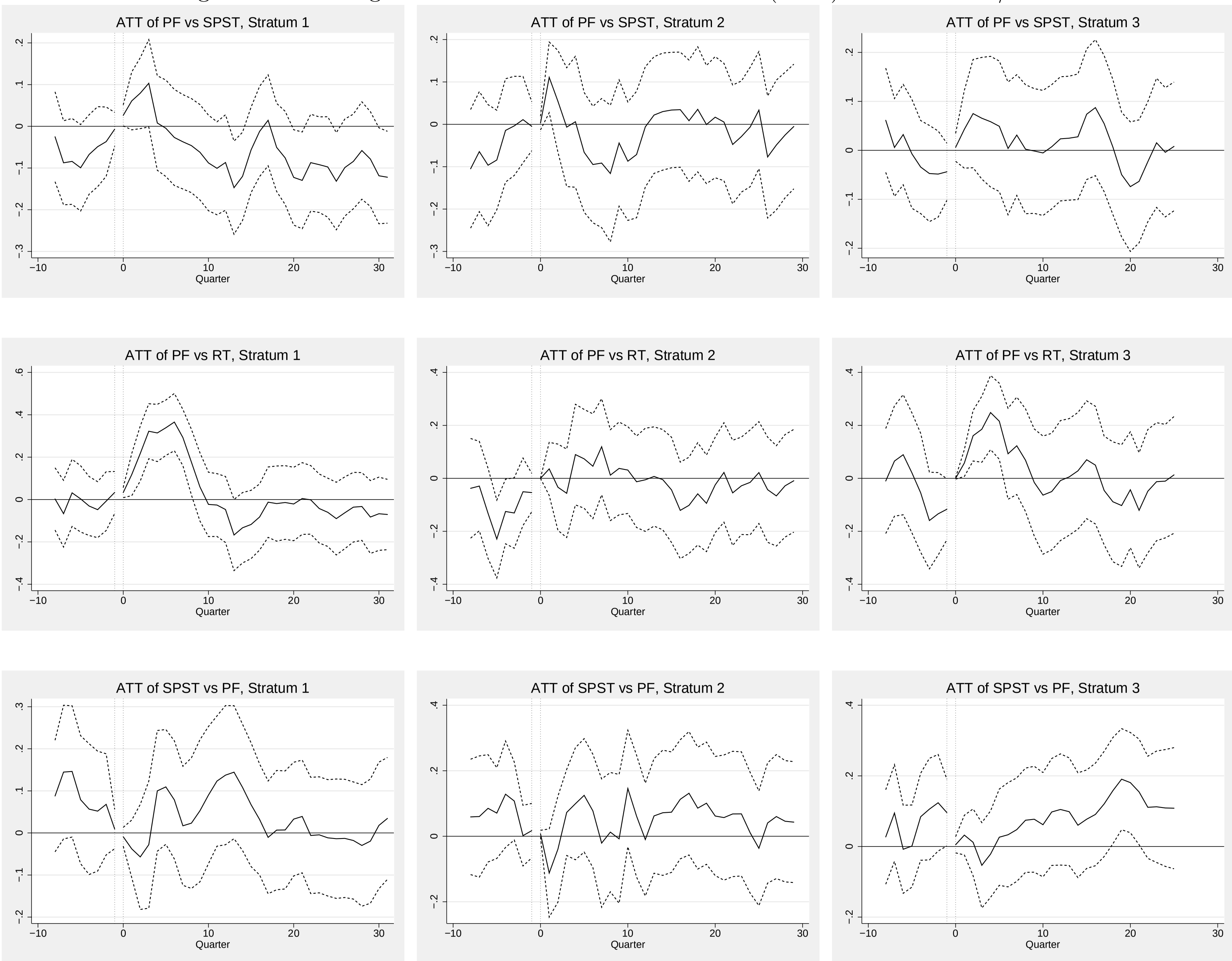
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) quarters on the abscissa.

Figure 1.5: Average Treatment Effect on the Treated (ATT) for Cohort 93/94



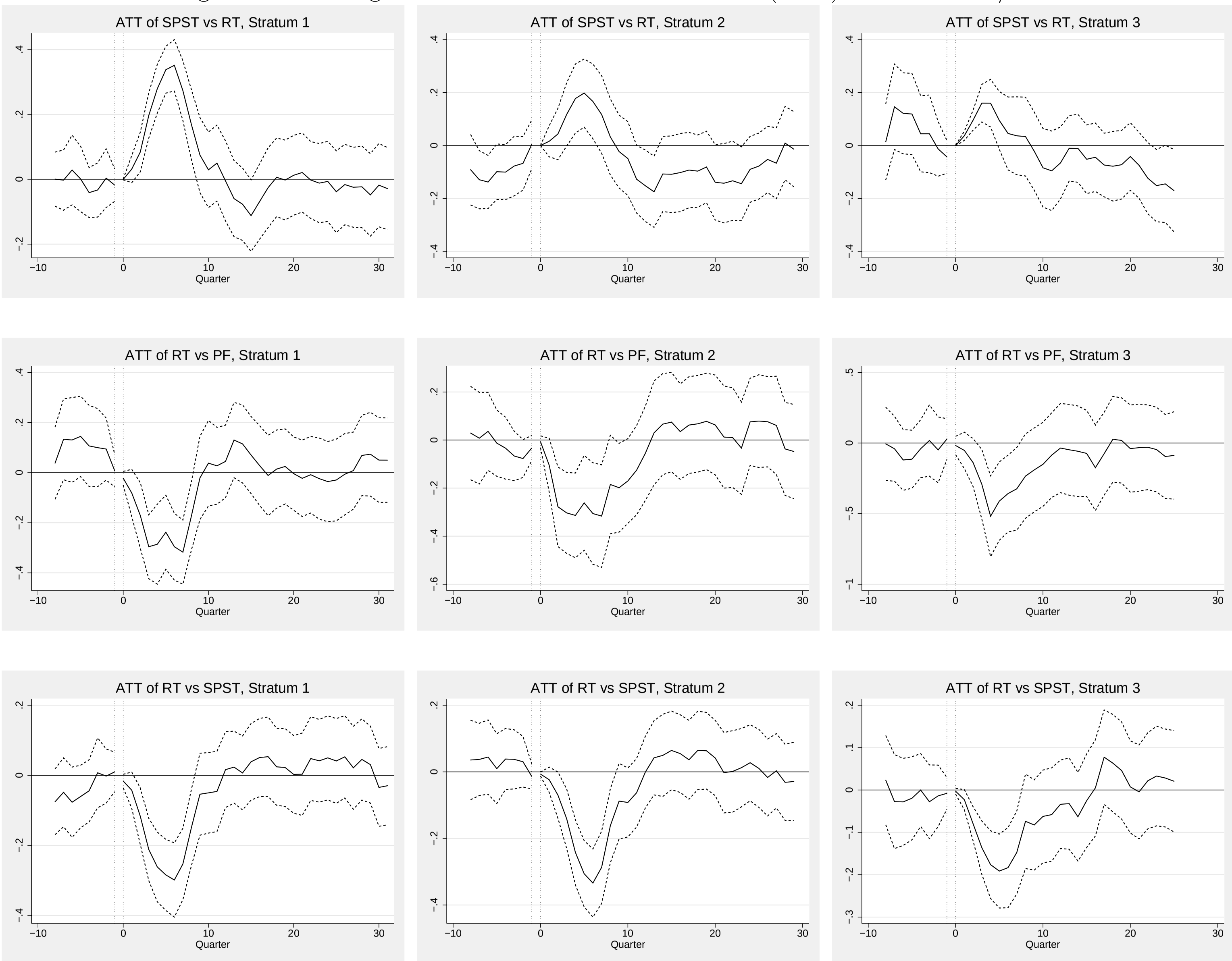
Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) quarters on the abscissa.

Figure 1.6: Average Treatment Effect on the Treated (ATT) for Cohort 93/94



Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) quarters on the abscissa.

Figure 1.7: Average Treatment Effect on the Treated (ATT) for Cohort 93/94



Difference in employment rates is measured on the ordinate, pre-unemployment (< 0) and post-treatment (≥ 0) quarters on the abscissa.

Table 1.7: Cumulated differences in employment rates – sum of quarter specific average treatment effects on the treated since beginning of treatment – Training versus Waiting

Cumulated Treatment Effects

Cohort 86/87

	8 quarters	16 quarters	24 quarters
	PF vs Waiting		
Stratum 1	-0.159 (0.382)	0.586 (0.706)	1.817 (1.018)*
Stratum 2	0.164 (0.316)	1.150 (0.653)*	1.971 (1.009)*
Stratum 3	0.276 (0.304)	0.748 (0.685)	1.280 (1.115)
	SPST vs Waiting		
Stratum 1	0.174 (0.118)	1.420 (0.241)***	2.524 (0.373)***
Stratum 2	0.631 (0.173)***	1.920 (0.353)***	2.766 (0.536)***
Stratum 3	0.702 (0.173)***	2.725 (0.406)***	4.221 (0.649)***
	RT vs Waiting		
Stratum 1	-1.353 (0.169)***	-0.150 (0.326)	0.921 (0.511)*
Stratum 2	-0.678 (0.252)***	1.069 (0.501)**	2.842 (0.761)***
Stratum 3	-0.347 (0.216)	1.673 (0.533)***	3.017 (0.808)***

Cohort 93/94

	8 quarters	16 quarters	24 quarters
	PF vs Waiting		
Stratum 1	-0.001 (0.293)	0.317 (0.606)	0.876 (0.924)
Stratum 2	0.340 (0.235)	1.566 (0.499)***	2.862 (0.744)***
Stratum 3	0.544 (0.276)**	1.590 (0.600)***	2.540 (0.899)***
	SPST vs Waiting		
Stratum 1	-0.012 (0.113)	1.201 (0.235)***	2.375 (0.348)***
Stratum 2	0.378 (0.130)***	1.745 (0.266)***	3.070 (0.421)***
Stratum 3	0.439 (0.097)***	1.495 (0.217)***	2.544 (0.338)***
	RT vs Waiting		
Stratum 1	-1.982 (0.149)***	-1.552 (0.340)***	-1.061 (0.535)**
Stratum 2	-1.218 (0.192)***	-0.059 (0.395)	1.352 (0.649)**
Stratum 3	-0.878 (0.260)***	-0.152 (0.563)	1.258 (0.904)

Remark: *, **, and *** denote significance at the 10%–, 5%–, and 1%–significance level, respectively.

Table 1.8: Averages of quarter specific average treatment effects on the treated

Average ATT

Cohort 86/87

	1st year	2nd year	3rd year	year 4 onwards
		Average ATT, PF vs Waiting		
Stratum 1	-0.086 (0.044)*	0.046 (0.061)	0.076 (0.057)	0.145 (0.048)***
Stratum 2	-0.039 (0.031)	0.080 (0.054)	0.115 (0.056)**	0.103 (0.052)**
Stratum 3	-0.041 (0.032)	0.110 (0.055)**	0.078 (0.059)	0.056 (0.059)
		Average ATT, SPST vs Waiting		
Stratum 1	-0.074 (0.013)***	0.118 (0.020)***	0.145 (0.020)***	0.143 (0.019)***
Stratum 2	-0.026 (0.018)	0.184 (0.030)***	0.179 (0.031)***	0.114 (0.027)***
Stratum 3	-0.003 (0.019)	0.178 (0.030)***	0.257 (0.034)***	0.198 (0.034)***
		Average ATT, RT vs Waiting		
Stratum 1	-0.221 (0.020)***	-0.117 (0.029)***	0.120 (0.029)***	0.143 (0.029)***
Stratum 2	-0.167 (0.024)***	-0.002 (0.046)	0.198 (0.043)***	0.229 (0.041)***
Stratum 3	-0.097 (0.020)***	0.010 (0.043)	0.206 (0.053)***	0.199 (0.049)***

Cohort 93/94

	1st year	2nd year	3rd year	year 4 onwards
		Average ATT, PF vs Waiting		
Stratum 1	-0.066 (0.032)**	0.066(0.047)	0.063 (0.047)	0.050 (0.044)
Stratum 2	-0.015 (0.026)	0.100 (0.044)**	0.148 (0.045)***	0.163 (0.040)***
Stratum 3	0.016 (0.028)	0.120 (0.047)**	0.113 (0.048)**	0.131 (0.045)***
		Average ATT, SPST vs Waiting		
Stratum 1	-0.104 (0.012)***	0.101 (0.020)***	0.156 (0.020)***	0.146 (0.017)***
Stratum 2	-0.047 (0.013)***	0.142 (0.022)***	0.177 (0.021)***	0.164 (0.022)***
Stratum 3	-0.015 (0.009)	0.124 (0.018)***	0.141 (0.019)***	0.127 (0.017)***
		Average ATT, RT vs Waiting		
Stratum 1	-0.255 (0.014)***	-0.241 (0.026)***	0.019 (0.033)	0.069 (0.030)**
Stratum 2	-0.162 (0.017)***	-0.143 (0.035)***	0.097 (0.036)***	0.179 (0.037)***
Stratum 3	-0.122 (0.031)***	-0.097 (0.044)**	0.072 (0.052)	0.154 (0.050)***

Remark: *, **, and *** denote significance at the 10%–, 5%–, and 1%–significance level, respectively.

Table 1.9: Cumulated differences in employment rates – sum of quarter specific average treatment effects on the treated since beginning of treatment – Pairwise comparisons of training programs

	Cumulated Treatment Effects		
Cohort 86/87			
	8 quarters	16 quarters	24 quarters
	PF vs SPST		
Stratum 1	0.028 (0.355)	-0.199 (0.686)	0.023 (1.036)
Stratum 2	-0.159 (0.426)	-0.014 (0.833)	0.431 (1.224)
Stratum 3	0.053 (0.331)	-1.016 (0.866)	-1.065 (1.455)
	PF vs RT		
Stratum 1	0.853 (0.395)**	0.348 (0.736)	0.259 (1.117)
Stratum 2	-0.650 (0.579)	-1.896 (1.101)*	-2.819 (1.726)
Stratum 3	1.237 (0.350)***	0.907 (0.836)	0.140 (1.402)
	SPST vs PF		
Stratum 1	-0.125 (0.339)	-0.848 (0.714)	-2.114 (1.041)**
Stratum 2	0 (0.437)	-0.380 (0.888)	-1.275 (1.375)
Stratum 3	0.798 (0.406)**	1.837 (1.022)*	1.768 (1.601)
	SPST vs RT		
Stratum 1	1.246 (0.354)***	1.072 (0.599)*	0.199 (0.789)
Stratum 2	0.380 (0.394)	-0.308 (0.761)	-1.177 (1.147)
Stratum 3	1.310 (0.286)***	1.625 (0.771)**	1.575 (1.240)
	RT vs PF		
Stratum 1	-1.121 (0.509)**	-0.309 (1.154)	0.523 (1.857)
Stratum 2	-0.496 (0.498)	0.413 (1.022)	1.252 (1.528)
Stratum 3	-0.133 (0.431)	1.498 (1.104)	1.632 (1.654)
	RT vs SPST		
Stratum 1	-1.173 (0.227)***	-1.024 (0.440)**	-0.774 (0.698)
Stratum 2	-0.868 (0.347)**	-0.142 (0.824)	1.225 (1.326)
Stratum 3	-0.430 (0.269)	-0.207 (0.691)	-0.066 (1.098)

Cumulated Treatment Effects – continued

Cohort 93/94

	8 quarters	16 quarters	24 quarters
		PF vs SPST	
Stratum 1	0.209 (0.282)	-0.498 (0.605)	-1.054 (0.930)
Stratum 2	-0.085 (0.354)	-0.324 (0.741)	-0.300 (1.136)
Stratum 3	0.333 (0.376)	0.485 (0.782)	0.439 (1.165)
		PF vs RT	
	8 quarters	16 quarters	24 quarters
Stratum 1	2.002 (0.376)***	1.723 (0.763)**	1.534 (1.234)
Stratum 2	0.274 (0.521)	0.295 (0.881)	-0.166 (1.285)
Stratum 3	1.084 (0.327)***	1.119 (0.971)	0.707 (1.616)
		SPST vs PF	
Stratum 1	0.174 (0.391)	0.920 (0.824)	1.017 (1.240)
Stratum 2	0.210 (0.366)	0.620 (0.828)	1.306 (1.374)
Stratum 3	0.081 (0.370)	0.733 (0.898)	1.852 (1.378)
		SPST vs RT	
Stratum 1	1.554 (0.201)***	1.623 (0.531)***	1.552 (0.905)*
Stratum 2	0.836 (0.339)**	0.126 (0.659)	-0.805 (1.002)
Stratum 3	0.629 (0.221)***	0.323 (0.570)	-0.337 (0.864)
		RT vs PF	
Stratum 1	-1.707 (0.374)***	-1.477 (0.805)*	-1.481 (1.164)
Stratum 2	-1.890 (0.445)***	-2.453 (1.017)**	-2.158 (1.678)
Stratum 3	-2.112 (0.743)***	-2.988 (1.713)*	-3.341 (2.694)
		RT vs SPST	
Stratum 1	-1.485 (0.257)***	-1.698 (0.540)***	-1.453 (0.848)*
Stratum 2	-1.411 (0.250)***	-1.661 (0.536)***	-1.389 (0.869)
Stratum 3	-0.940 (0.201)***	-1.372 (0.519)***	-1.122 (0.825)

Remark: *, **, and *** denote significance at the 10%–, 5%–, and 1%–significance level, respectively.

Chapter 2

Long–Run Effects of Training Programs for the Unemployed in East Germany

2.1 Introduction

Active labor market policy (ALMP) has been used at an unprecedently high scale during the transition process in East Germany in the 1990s. Public sector sponsored training has been a major part of ALMP with the goal to adjust the skills of the East German workforce to the needs of a Western market economy. Annual entries into training programs were around 250 thousand during the years 1993 to 1996 (Bundesanstalt für Arbeit 1993, 1997, 2001). In comparison to public sector sponsored training in other countries, the East German experience shows the following five specific aspects. First, participants had fairly high levels of formal education. Second, access to treatment was easy since targeting was very low. Third, the market for training provision had to be established and in the early 1990s case workers had no practical experience on what works. Fourth, predictions about the catching up process of East Germany and about future labor market trends proved to be wrong. Fifth, the duration of training programs was fairly long.

During the last decade, there were a lot of pessimistic assessments regarding the usefulness of public sector sponsored training programs in raising employment chances of the unemployed (see the surveys in Fay, 1996; Heckman et al., 1999; Martin and Grubb, 2001; Kluve and Schmidt, 2002). These studies doubt that large scale training programs, which are not well targeted, are successful in raising employment. However, evidence for Eastern European transition economies (other than East Germany) has often shown positive effects (Kluve et al., 2004; Lubyova and van Ours, 1999; Puhani, 1999). Recently, OECD (2005) has argued that long-term labor market programs, such as training, often have little or negative short–run effects on outcomes, which can be attributed to lock–in effects. However, in some cases, positive long–term effects exist for long training programs, for which lock–in effects are worse than for short programs (see also Fay, 1996). Therefore, it is crucial to assess program impacts in a longer term perspective in order to investigate whether the sizeable lock–in effects in the short run are compensated by positive long run effects.

For East Germany, appropriate data for a long term evaluation of public sector sponsored training were not available for a long time and, until recently, the available evidence has been quite mixed.[1] Detailed administrative data have been used in

[1]See Bergemann et al. (2004), Fitzenberger and Prey (2000), Kraus et al. (1999), or Lechner (2000) for exemplary studies based on survey data. Speckesser (2004, chapter 1) and Wunsch

the recent studies of Lechner et al. (2005b) (≡LMW), Fitzenberger and Speckesser (2007) (≡FS), and Hujer et al. (2006) (≡HTZ), where the first two studies are based on the same data as this study, while the third uses administrative data since 2000. HTZ find negative short–run effects which are probably driven by lock–in effects, while their data do not allow to investigate long–run effects. LMW and FS find positive medium– and long–run employment effects for some treatments considered in this paper. LMW evaluate effects of three training programs (long training, short training, retraining) on employment and benefit recipiency. They find strong evidence that, on average, the training programs under investigation increase long–term employment prospects and do not change benefit recipiency. As important exceptions, long training and retraining show no positive employment effects for males. FS estimate the employment effects of one major training program (Provision of Specific Professional Skills and Techniques, SPST) against nonparticipation in SPST for 36 months after the beginning of the treatment. The analysis is performed only for the 1993 inflow sample into unemployment. The analysis finds positive medium–run employment effects, but it does not distinguish between genders.

The vast majority of the existing evaluation studies for East Germany uses a static evaluation approach, which contrasts receiving treatment during a certain period of time against the alternative of not receiving treatment during this period of time (FS and HTZ are recent exceptions). In a dynamic setting, the timing of events becomes important, see Abbring and van den Berg (2003), Fredriksson and Johansson (2003, 2004), and Sianesi (2003, 2004). Static treatment evaluations run the risk of conditioning on future outcomes, leading to possibly biased treatment effects. This paper follows Sianesi (2003, 2004) and estimates the effects of treatment starting after some unemployment experience against the alternative of not starting treatment at this point of time and waiting longer. The actual implementation of the estimator builds and extends upon FS and Fitzenberger et al. (2006). The estimated dynamic treatment effects mirror the decision problem of the case worker and the unemployed who decide recurrently during the unemployment spell, whether to begin any program now or to postpone participation to the future.

Using a dynamic multiple treatment framework, this study analyzes the effects of three exclusive training programs (practice firms, SPST, retraining) for inflow samples into unemployment for the two years 1993/94. We evaluate medium– and

(2006, section 6.5) provide comprehensive surveys of this literature, which is not reviewed here for the sake of brevity, and discuss critically the data used.

long–run treatment effects both for employment and benefit recipiency up to 24–30 quarters after the beginning of the treatment depending on the starting date of the treatment. The analysis is performed separately for males and females to reexamine the evidence in LMW and the two studies differ substantially regarding the exact treatment definition, the choice of valid observations, and the econometric methods used. Our results confirm the positive employment effects for SPST reported in FS after the initial negative lock–in effect to hold for a much longer time period and to apply for both males and females. Our study finds no positive employment effects for practice firms and in four out of six cases for retraining. We do not find systematic gender differences and, similar to the study Fitzenberger et al. (2006) for West Germany, our assessment of retraining is considerably worse than of SPST. Furthermore, we do not find any of the three programs to reduce significantly the benefit recipiency rate in the medium and long run. In the short run, all programs show the lock–in effect with an increase in the benefit recipiency rate, thus providing evidence for 'benefit churning' as in Kluve et al. (2004).

Our analysis differs considerably from the recent work of LMW, FS, and HTZ. LMW use a static multiple treatment evaluation approach. They find gender differences for long training and retraining, which we can not replicate using our dynamic evaluation approach. We explore potential reasons for the different results. LMW analyze the effects of treatments starting during the years 93/94 for unemployed whose unemployment spells start during the years 93/94. We also analyze the inflows into unemployment for the years 93/94 but we analyze the effects of all treatments taking place during the first two years of unemployment. We investigate whether the estimated treatment effects differ for treatments during three different time windows of elapsed unemployment durations. Furthermore, there are a number of important differences in the definition of the treatments, the selection of samples, and the implemented methods. FS use a dynamic treatment evaluation approach for SPST only. We estimate the effects of three training programs for a much longer time period after the beginning of the program, and our analysis distinguishes between genders. In addition to the employment effects, we also analyze the effects on benefit recipiency. Furthermore, we use a larger inflow sample than FS. In contrast to HTZ, who estimate a duration model and focus on exits from unemployment, we estimate medium– and long–run effects on both employment and benefit recipiency, which we distinguish from lock–in effects. Estimating a duration model, it would be very difficult to take account of the large number of exits into and out of employment observed after the first exit from unemployment.

The remainder of this paper is structured as follows: Section 2.2 gives a short description of the institutional regulation and participation figures for active labor market policy. Section 2.3 focuses on the different options of further training, their target groups, and course contents. Section 2.4 describes the methodological approach to estimate the treatment effects. The empirical results are discussed in section 2.5. Section 2.6 concludes. The final appendix provides further information on the data and detailed empirical results. An additional appendix includes further details on the data and the empirical results.

2.2 Basic Regulation and Programs

2.2.1 Basic Regulation

For the time period considered here, public sector sponsored training in Germany is regulated by the Labor Promotion Act (*Arbeitsförderungsgesetz*, AFG) and is offered and coordinated by the German Federal Employment Office (formerly *Bundesanstalt für Arbeit*, BA). We consider the two main training programs: **Further training** (*Weiterbildung*) includes the assessment, maintenance and extension of skills, including technical development and career advancement. The duration of the courses depends on individual predispositions and adequate courses provided by the training suppliers. **Retraining** (*Umschulung*) enables vocational re–orientation if a completed vocational training does not lead to adequate employment. Retraining is supported for a period up to 2 years and aims at providing a new certified vocational education degree.

2.2.2 Evaluated Programs

Further training is a very broad legal category and consists of quite heterogeneous programs. Hence we utilize a classification developed in FS and evaluate two specific further training programs: Practice Firms (PF) and provision of specific professional skills and techniques (SPST).

Practice Firms (PF) are simulated firms in which participants practice everyday working activities. The areas of practice are whole fields of profession, not specific professions. Hence, practice firms mainly train general skills while provision of new professional skills is of less importance. Some of the practice firms are technically oriented, the practice studios, whereas others are commercially oriented, the practice enterprises. One of the practice firm's goals is to evaluate the participant's aptitude for a field of profession. The programs usually last for six months and do not provide official certificates.

Provision of **Specific Professional Skills and Techniques (SPST)** intends to improve the starting position for finding a new job by providing additional skills and specific professional knowledge in medium–term courses. It involves refreshing specific skills, e.g. computer skills, or training on new operational practices. SPST mainly consists of classroom training but an acquisition of professional knowledge through practical work experience may also be provided. After successfully completing the course, participants usually obtain a certificate indicating the contents of the course, i.e. the refreshed or newly acquired skills and the amount of theory and practical work experience. Such a certificate is supposed to serve as an additional signal to potential employers and to increase the matching probability since the provision of up to date skills and techniques is considered to be a strong signal in the search process. The provision of specific professional skills and techniques aims at sustained reintegration into the labor market by improving skills as well as providing signals.

Compared to retraining, which is a far more formal and thorough training on a range of professional skills and which provides a complete vocational training degree, the role of SPST for a participant's occupational knowledge is weaker. However, the amount of occupation specific knowledge imparted in SPST certainly exceeds the level provided in short–term programs (not evaluated here) that usually aim at improving job search techniques or general social skills. Thus, SPST ranges in the middle between very formal (and very expensive) courses and very informal and short courses (improving general human capital).

Retraining (RT) consists of the provision of a new and comprehensive vocational training according to the regulation of the German apprenticeship system. It is targeted to individuals who already completed a first vocational training and face severe difficulties in finding a new employment within their profession. It might however also be offered to individuals without a first formal training degree if they fulfill additional eligibility criteria.

Retraining provides widely accepted formal certificates. It comprises both, theoretical training and practical work experience. The theoretical part of the formation takes place in the public education system. The practical part is often carried out in firms that provide work experience in a specific field to the participants, but sometimes also in interplant training establishments. This type of treatment leads to a certified job qualification in order to improve the job match. Ideally, the training occupation in retraining corresponds to qualifications which are in high demand in the labor market.

2.2.3 Financial Incentives for Participation

Participants in the training programs considered are granted an income maintenance (IM, *Unterhaltsgeld*). To qualify, they must have been employed for at least one year or they must be entitled to unemployment benefits or subsequent unemployment assistance.[2]

Since 1994, IM is equal to the standard unemployment benefits (UB, *Arbeitslosengeld*). It amounts to 67% of previous net earnings for participants with at least one dependent child and 60% otherwise (note that in 1993 replacement ratios for IM were higher at 73% and 65%, respectively). In contrast, unemployed, whose UB expired, can receive the lower, means tested unemployment assistance (UA, *Arbeitslosenhilfe*) which amounts to 57% (with children) and 53% (without children). This means that for these unemployed IM during the program is higher than UA. Additionally, participants could defer the transition from UB to the lower UA and, in some cases, even requalify for the higher UB.

Concluding, there are positive financial incentives for the unemployed to join a program. In addition, the BA bears all costs directly incurred through participation in a further training scheme, especially course fees.

2.3 Data

We use a database which integrates administrative individual data from three different sources (see Bender et al. (2005) for a detailed description). The data contain

[2]For a more detailed description of the institutions, see Bender et al. (2005), Fitzenberger, Osikominu and Völter (2006), or Wunsch (2006).

spells on

- employment subject to social insurance contributions,
- transfer payments by the BA,
- and participation in training programs.

Further details on the compilation of the data can be found in the additional appendix.

The basic data source is the **IAB Employment Subsample** (*IAB Beschäftigtenstichprobe*, IABS) for the time period 1975–97, see Bender et al. (2000) and Bender et al. (2005, chapter 2.1). The IABS is a 1% random sample drawn from employment register data for all employees subject to social insurance contributions. Therefore, we restrict the analysis to inflows from employment to unemployment. For this study, we merge additional information for 1998–2002 to the basic data.

The second important source is the **Benefit Payment Register** (*Leistungsempfängerdatei*, LED) of the Federal Employment Office (BA), see Bender et al. (2005, chapter 2.2). These data consist of spells on periods of transfer payments granted by the BA to unemployed and program participants. Besides unemployment benefit or assistance, these data also record very detailed information about income maintenance payments related to the participation in training programs.

The third data source records training participation (**FuU-data**). The BA collects these data for all participants in further training, retraining, and other training programs for internal monitoring and statistical purposes, see Bender et al. (2005, chapter 2.3). For every participant the FuU-data contains detailed information about the program and about the participant.

The FuU–data were merged with the combined IABS–LED data by social insurance number and additional covariates. Numerous corrections have been implemented in order to improve the quality of the data, see Bender et al. (2005, chapters 3–4), FS, and the additional appendix for more information. The IABS provides information on personal characteristics and employment histories. The combination of the transfer payment information and the participation information is used to identify the likely participation status regarding the different types of training programs.

When an individual is not observed in any of the three spell types (employment, transfers, training participation), we interpret this as being out of the labor force. The spell information on the employment state of an individual is first transformed into monthly dummy variables (based on the dominating state). We construct separately monthly dummy variables for training status. Then, for our analysis, the data is aggregated to a quarterly frequency.

Inflow Sample into Unemployment: To analyze the effect of training programs on employment and benefit recipiency of unemployed individuals, we base our empirical analysis on the sample of inflows into unemployment during the years 1993/94 in East Germany, omitting Berlin. We consider individuals who experience a transition from employment to nonemployment and for whom a spell with transfer payments from the Federal Employment Office starts during the first 12 months of nonemployment or for whom the training data indicate program participation before a new job is found.[3] The start of the nonemployment spell is denoted as the beginning of the unemployment spell. We condition on receipt of unemployment compensation or program participation to exclude individuals who move out of the labor force.[4] This rule concerns almost exclusively individuals who do not participate in any training program during their nonemployment spell. A treatment is only considered if the unemployed does not start employment before the second month of treatment (to omit training while holding a job). Furthermore, we restrict our samples to the 25 to 55 years old in order to rule out periods of formal education or vocational training as well as early retirement. For RT, we restrict the sample to the 25 to 50 years old.

We choose the years 1993/94 because data for East Germany start in 1992 and we want to control for one year of labor market experience before the beginning of unemployment. Our merged data allow to follow individuals until the end of 2002. Table 2.1 gives information about the size of the inflow samples and the incidence of training.

Participation by Type of Training: We focus on the three types of training programs PF, SPST, and RT, as described in section 2.2.2 above. These programs are targeted to the unemployed and do not involve on–the–job training (training while working in a regular job). The total inflow sample comprises 6,135 spells for women and 5,911 spells for men. There are 1,550 training spells for females and 835

[3]This design allows the same individual to be in the sample more than once if it has more than one transition from employment to unemployment in 1993/94.

[4]Only 1% of training participants do not receive transfer payments during the first 12 months.

for men. Thus, about 25% of the females and 14% of the men participate in one of the three training programs considered, which reflects the large scale of training programs during the East German transition process. Among these programs, SPST represents the largest with 78% and 63% of the training spells, respectively for females and males. For females 13% and for men 28% of all training spells are RT, and PF represents the smallest group in both samples. In absolute numbers, there are 145 (73) PF spells in the female (male) inflow sample, 1,210 (528) SPST spells and 195 (234) RT spells. Table 2.2 shows the frequency of training by elapsed duration of unemployment.

Table 2.3 provides descriptive statistics on the elapsed duration of unemployment at the beginning of treatment. Our discussion focuses on quantiles because averages may be misleading. The median entrant in PF has been unemployed for 10 months for females and only for 5 months for males. Late starts (75%–quantile) of PF occur after 14 months for females and after 11 months for men. SPST is the program which starts latest with a median of 11 months for females and 7.5 months for men. RT is the program which starts the earliest for females. The median is 8 months. The median for males of 6 months is higher than the value for PF. In general, females start later than men.

Table 2.4 provides descriptive information on the duration of training spells. The average durations are quite different between the programs but comparable across genders. Participation in PF is shortest. On average woman stay 6.5 months in PF and men 6.1 months. Participation in SPST has an average duration of 9.1 months for females and 8.8 months for males. Participation in RT lasts almost twice as long as in SPST with an average of 18.7 months for women and 17.3 months for men.

2.4 Evaluation Approach

Our goal is to analyze the effect of $K = 3$ different training programs on two outcome variables, namely the individual quarterly employment rate (ER) and the individual quarterly benefit recipiency rate (BR), both measured as quarterly averages of monthly dummy variables.[5] In a situation where individuals have multiple treatment options, we estimate the average treatment effect on the treated (ATT)

[5]These quarterly rates can take the four values 0, 1/3, 2/3, and 1.

of one training program against nonparticipation in any of the three programs.[6] Extending the static multiple treatment approach to a dynamic setting, we follow Sianesi (2003, 2004) and apply the standard static treatment approach recursively depending on the elapsed unemployment duration. This dynamic evaluation approach is implemented for our problem as in FS and Fitzenberger et al. (2006). The estimated dynamic ATT parameters mirror the decision problem of the case worker and the unemployed who decide recurrently during the unemployment spell, whether to begin any program now or to postpone participation to the future.

Our empirical analysis is based upon the potential–outcome–approach to causality, see Roy (1951), Rubin (1974), and the survey of Heckman, LaLonde, Smith (1999). Lechner (2001) and Imbens (2000) extend this framework to allow for multiple, exclusive treatments. Let the 4 potential outcomes be $\{Y^0, Y^1, Y^2, Y^3\}$, where $Y^k, k = 1, ..., 3$, represents the outcome associated with training program k and Y^0 is the outcome when participating in none of the 3 training programs. For each individual, only one of the $K + 1$ potential outcomes is observed and the remaining K outcomes are counterfactual. We estimate the average treatment effect on the treated (ATT) of participating in treatment $k = 1, 2, 3$ against nonparticipation $k = 0$.[7]

Fredriksson and Johansson (2003, 2004) argue that a static evaluation analysis, which assigns unemployed individuals to a treatment group and a nontreatment group based on the treatment information observed in the data, yields biased treatment effects. This is because the definition of the control group conditions on future outcomes or future treatment. For Sweden, Sianesi (2004) argues that all unemployed individuals are potential future participants in active labor market programs, a view which is particularly plausible for countries with comprehensive systems of active labor market policies (like Germany).[8] This discussion implies that a purely

[6] For the large scale training programs implemented in East Germany, it is likely that the Stable–Unit–Treatment–Value assumption is violated because general equilibrium effects exist. What we actually estimate is a meaningful ATT for the following scenario: Suppose one changes treatment status for one random treated individual to being nontreated (including the possibility of later treatment). The estimated ATT estimates the expected value of the change in the outcome variable times -1. In this sense, our estimates reflect meaningful effects of such a marginal policy change.

[7] Using the same approach, a pairwise comparison of the differential effects of the programs would be feasible, see Lechner (2001) or Fitzenberger et al. (2006). Such a pairwise comparison is not pursued in this paper for the sake of space.

[8] In East Germany, active labor market programs were implemented after unification at an unprecedented scale.

static evaluation of the different training programs is not warranted. Following Sianesi (2003, 2004), we analyze the effects of the first participation in a training program during the unemployment spell considered *conditional on the starting date of the treatment*. We distinguish between treatment starting during quarters 1 to 2 of the unemployment spell (stratum 1), treatment starting during quarters 3 to 4 (stratum 2), and treatment starting during quarters 5 to 8 (stratum 3).

Our approach differs from Sianesi (2003, 2004) and Fredriksson and Johansson (2003, 2004) in the following important aspects. We believe the starting date of the treatment to be somewhat random (relative to the elapsed duration of the unemployment spell) due to available programs starting only at certain calendar dates. Based on this argument, we pool the treatment Probit for all eligible persons in unemployment for a longer stratum than Sianesi (2003, 2004), who uses months, and we exclude individuals who start treatment later within the same stratum as possible control persons. Fredriksson and Johansson (2003, 2004) define a treatment parameter which integrates the treatment effects over program starting dates. In devising their estimators in discrete time, these authors assume that all individuals who do not receive treatment during the next (short) time period can serve as control persons for estimating counterfactual hazard rates. We do not think that this assumption is justified in our context because of the aforementioned randomness of starting dates.

Our estimated ATT parameter has to be interpreted in a dynamic context. We analyze treatment conditional upon the unemployment spell lasting at least until the start of the treatment k and this being the first treatment during the unemployment spell considered. Therefore, the estimated treatment parameter is

$$\begin{aligned}(2.1)\qquad \theta(k;u,\tau) &= E(Y^k(u,\tau)|T_u = k, U \geq u-1, T_1 = ... = T_{u-1} = 0)\\ &\quad -E(Y^0(u,\tau)|T_u = k, U \geq u-1, T_1 = ... = T_{u-1} = 0)\ ,\end{aligned}$$

where T_u is the treatment variable for treatment starting in quarter u of unemployment and U is the completed duration of the unemployment spell. $Y^k(u,\tau)$ and $Y^0(u,\tau)$ are the potential treatment outcomes for treatments k and 0, respectively, in periods $u+\tau$, where treatment starts in period u and $\tau = 0,1,2,...$, counts the quarters since the beginning of treatment. The nontreatment outcome $Y^0(u,\tau)$ refers to the case where the individual does not receive any treatment until the end of the stratum considered. Actually, we estimate the treatment parameter

$\theta(k;\tau) = \sum_u g_u \theta(k;u,\tau)$, which is averaged within a stratum with respect to the distribution g_u of starting dates u.

We evaluate the differential effects of multiple treatments assuming the following dynamic version of the conditional mean independence assumption (DCIA)[9]

$$E(Y^0(u,\tau)|U \geq u-1, T_1 = ... = T_{u-1} = 0, T_u = k, X, ben(u)) \tag{2.2}$$

$$= E(Y^0(u,\tau)|U \geq u-1, T_1 = ... = T_{u-1} = T_u = ... = T_{\bar{u}} = 0, X, ben(u)) \ ,$$

where X are time–invariant (during the unemployment spell) characteristics, $ben(u)$ is the number of months the unemployed were receiving benefits during the unemployment spell before the start of the treatment u, and $\bar{u}$ denotes the last quarter of the stratum considered. We effectively assume that conditional on X, conditional on being unemployed until period $u-1$, conditional on having received benefits the same number of months before u, and conditional on not having received a treatment before u, individuals treated in u are comparable in their nontreatment outcome to individuals who do not start any treatment until $\bar{u}$ (recall from above, that $Y^0(u,\tau)$ involves no treatment until $\bar{u}$).

Building on Rosenbaum and Rubin's (1983) result on the balancing property of the propensity score in the case of a binary treatment, Lechner (2001) shows that the conditional probability of treatment k, given that the individual receives treatment k or no treatment 0, $P^{k|k0}(X)$, exhibits an analogous balancing property for the pairwise estimation of the ATT's of program k versus no participation 0. This allows to apply standard binary propensity score matching based on the sample of individuals participating in either program k or in no program 0 (Lechner, 2001; Gerfin and Lechner, 2002; Sianesi, 2003). For this subsample, we simply estimate the probability of treatment k and then apply a bivariate extension of standard propensity matching techniques. Implicitly, we assume that the actual beginning of treatment within a stratum is random conditional on X.

To account for the dynamic treatment assignment, we estimate the probability of treatment k given that unemployment lasts long enough to make an individual 'eligible'. For treatment during quarters 1 to 2, we take the total sample of unemployed,

[9] In addition to DCIA, we also assume that the probability of treatment is less than one conditional on the conditioning variables in equation (2.2) and that the Stable Unit Treatment Value assumption holds. These are further assumptions needed to estimate an ATT parameter, see Heckman, LaLonde, Smith (1999).

who participate in k or in no program during quarters 1 to 2 (stratum 1), and estimate a Probit model for participation in k. This group includes those unemployed who either never participate in any program or who start some treatment after quarter 2. For treatment during strata 2 and 3, the basic sample consists of those unemployed who are still unemployed at the beginning of the stratum.

We implement a stratified local linear matching approach by imposing that the matching partners for an individual receiving treatment k are still unemployed in the quarter (of elapsed unemployment duration) before treatment k starts and have received benefits the same number of months until the quarter before treatment starts. The expected counterfactual employment outcome for nonparticipation is obtained by means of a local linear regression on the propensity score and the starting month of the unemployment spell to match on calender time. We use a bivariate crossvalidation procedure to obtain the bandwidths in both dimensions (propensity score and beginning of unemployment spell).

An estimate for the variance of the estimated treatment effects is obtained through bootstrapping based on 200 resamples.[10] Effectively using a block bootstrap procedure for clustered inference, we resample the entire time series of observations for one individual. Our bootstrap procedure takes account of the sampling variability in the estimated propensity score and it is autocorrelation robust.

As a balancing test, we use the regression test suggested in Smith and Todd (2005) to investigate whether the time–invariant (during the unemployment spell) covariates are balanced sufficiently by matching on the estimated propensity score $P^{k|k0}(X)$ using a flexible polynomial approximation. Furthermore, we investigate whether treated and matched nontreated individuals differ significantly in their outcomes before the beginning of treatment, in addition to those already used as arguments of the propensity score. We estimate these differences in the same way as the treatment

[10] Abadie and Imbens (2006) show that the bootstrap fails for nearest neighbor matching because of a lack of smoothness resulting in local convergence not being uniform (see also Heckman et al., 1998, p. 276). In contrast, local linear matching with appropriate trimming to guarantee common support and under a weak convergence condition for the bandwidth parameters, is shown by Heckman et al. (1998, p. 278) to exhibit sufficiently smooth convergence for standard asymptotic distribution theory to hold. In particular, the estimated ATT parameter has a standard asymptotically linear representation and it is asymptotically normally distributed with $\sqrt{N}$ convergence rate. Although we are not aware of a formal proof, the bootstrap is therefore likely to be valid for local linear matching. Horowitz (2001, section 2) discusses the consistency of the bootstrap for $\sqrt{N}$ asymptotically normal estimators with an asymptotically linear representation. Although local linear matching involves an intermediate nonparametric estimation step, a similar result is likely to hold.

effects after the beginning of the program. By construction, treated individuals and their matched counterparts exhibit the same unemployment duration until the beginning of treatment.

Finally, we need to discuss why we think that the DCIA (2.2) is plausible for our application. As Sianesi (2004), we argue that the participation probability depends upon the variables determining re-employment prospects once unemployment began. Consequently, all individuals are considered who have left employment in the same two years (matching controls for beginning of unemployment) and who have experienced the same unemployment duration and the same number of months receiving benefits before program participation. Furthermore, observable individual characteristics and information from the previous employment spell have been included in the propensity score estimation. E.g., we consider skill information, regional information, occupational status, and industry which should be crucial for re-employment chances. Unfortunately, our data lack subjective assessments of labor market chances of the unemployed (e.g. by case workers). We argue that these are proxied sufficiently by the observed covariates in so far as they affect selection into the program. This is particularly plausible, since participation occurred at a large scale, assignment was not very targeted, and case workers lacked practical experience on 'what works' in a quickly changing economic environment. Supporting our point of view, Schneider et al. (2006) argue that until 2002 assignment to training was strongly driven by the supply of available courses.

2.5 Empirical Results

2.5.1 Estimation of Propensity Scores

Our empirical analysis is performed separately for females and males. To estimate the propensity scores, we run Probit regressions for each of the three programs for taking part in this program versus not taking part in any program ("waiting") for training starting during the three time intervals for elapsed unemployment duration, i.e. 1–2 quarters (stratum 1), 3–4 quarters (stratum 2), and 5–8 quarters (stratum 3). The additional appendix reports our preferred specifications, which are obtained after extensive specification search, summary statistics of the covariates used, detailed results of the balancing tests, and figures on common support.

The covariates considered are all defined for the beginning of unemployment and are thus time–invariant for an individual during the unemployment spell. Personal characteristics considered are age, marital status and formal education (with/without vocational training degree, tertiary education degree). In addition, we use information about the last employer, namely industrial sector and firm size, and a number of characteristics of the previous job such as employment status and information on earnings in the previous job. Regarding the employment and program participation history, we consider the employment history and participation in any ALMP program in the year before the beginning of unemployment. Differences in regional labor market conditions as well as supply of programs are the reason to include regional variables in the specification. We use the federal state and the population density at the district level. Finally, we also use the calendar month of the beginning of the unemployment period.

Our specification search starts by using as many as possible of the covariates mentioned above without interactions. The specification search is mainly led by the following two criteria: (i) single and joint significance, and (ii) balance of the covariates according to the regression based balancing test in Smith and Todd (2005). In general, insignificant covariates are dropped. We also test for the significance of interaction effects, in particular interactions with age. In order to achieve balance of covariates, we test different functional forms and interaction effects. In a few cases, we keep insignificant covariates or interactions, when they help to achieve balance. As we find the balancing test to be somewhat sensitive to small cell sizes we occasionally aggregate small groups that have similar coefficients.

The results for the Probit estimates show that the final specifications vary considerably between men and women and the three time intervals for a given program. Age effects are significant in most cases. In particular, participants in retraining are younger than individuals in other groups.

Our chosen specifications for the propensity score pass the regression based balancing test (no rejection) of Smith and Todd (2005) for a sufficiently large number of covariates. We graphically examine the common support requirement for estimating the average treatment effect on the treated (ATT). Overall, we are satisfied with the overlap of support in all cases and proceed without restricting the samples.[11]

[11] In four cases (out of 16) we have to drop one and in one case two treated individuals from the treatment effect estimations due to numerical problems.

2.5.2 Estimated Treatment Effects

We estimate the effects of the three types of training programs PF, SPST, and RT, separately for males and females. The two outcome variables considered are the individual quarterly employment rate (ER) and benefit recipiency rate (BR: UB, UA, or IM; see section 2.2.3). We match participants in treatment k and nonparticipants in any treatment, who are still unemployed in the quarter before treatment starts and have received benefits the same number of months until the quarter before treatment starts, by their similarity in the estimated propensity scores and the starting month of the unemployment spell. The ATT is then estimated separately for quarters τ since the beginning of program k for stratum 1, 2, and 3. In the case of PF for males, we only estimate the ATT for stratum 1, because the number of treated individuals in strata 2 and 3 is very small, see table 2.2.

Figures 2.1–2.6 display the estimated treatment effects $\hat{\theta}(k;\tau)$ on the vertical axis against quarter $\tau \geq 0$ since the beginning of treatment or quarter $\tau < 0$ before the beginning of treatment. The time axis is divided into three parts by two vertical lines, which denote the last quarter before the unemployment spell starts and the treatment start $\tau = 0$, respectively. The left part shows the four quarters before unemployment starts, the middle part the gap between the beginning of the unemployment spell and the beginning of treatment and the right part the time since treatment start. Each figure contains a panel of three times four graphs (except PF for males, with only stratum 1 in figure 2.2), where each row represents one stratum of elapsed duration of unemployment. The first and third column show the evolution of average outcomes for treated individuals (solid line) and their estimated nontreatment counterfactual (dashed line). The differences of these lines are displayed in the second and fourth column (solid line), respectively, as the estimated treatment effects together with pointwise 95%–confidence bands (dashed lines).

To summarize the graphical evidence in a systematic way, tables B.26 and 2.7 provide cumulated treatment effects ($\sum_{\tau=0}^{L-1} \hat{\theta}(k;\tau)$) over the first $L = 8$, 16, and 24 quarters since beginning of treatment and average treatment effects during quarter 4 to 23 and 8 to 23 $[1/(24-l)\sum_{\tau=l}^{23} \hat{\theta}(k;\tau)$ for $l = 4, 8]$.[12] These aggregated effects are calculated as sums or averages of the effects over time since treatment start τ. Note that we do not sum across strata k, i.e. we do not aggregate treatment starting dates

[12]For the bootstrap, we calculate the aggregate effect for each resample and then take the empirical standard deviation across the resamples as our estimate of the standard error of the estimated aggregate effect in the sample.

across strata – something which is done in Fredriksson and Johansson (2004). The cumulated effects as sums of the graphically depicted quarterly effects are measured in quarters of employment or benefit recipiency. They are a way of contrasting the initial negative lock-in-effects with the later positive program effects and can be seen as the net present value of participation measured in quarters of the outcome variable assuming a discount rate of zero. When the cumulated employment effect is positive the lower employment rate during the lock-in-effect can be compensated by the higher employment rate afterwards. And with benefit recipiency it is a negative cumulated effect, which indicates that a lower recipiency rate in the medium and long run can compensate for the higher rate during the lock-in. The average effects over the period between one or two years and six years from the beginning of treatment serve the purpose of giving precisely estimated medium– to long–run program effects unaffected by the short–run lock-in-effects.

The treatment PF (figures 2.1–2.2) basically shows significant (here and in the following, significant refers to statistically significant effects) negative lock–in effects on ER during the first six quarters (the solid line in the first columns lies below the dashed line)[13] and no significant positive ER effects afterwards. The BR effects are almost symmetric, with positive BR effects during the lock–in period and mostly no significant BR effects afterwards, except for stratum 3 for women where the BR effect seems to be quite volatile and often significantly positive in the medium– and long–run. The results are quite similar in stratum 1 for both genders. The graphical evidence is confirmed in tables B.26 and 2.7. We restrict our discussion of the aggregated effects to the cumulated effects over 24 quarters and to the average effects during quarter 8 to 23. None of the aggregated ER effects is significant. For BR, we find no significant aggregated effects on women for stratum 1 and 2. For Men in stratum 1 the cumulated effect on BR is significantly positive, but the average effect is insignificant. For stratum 3, we find both effects to be significantly positive. Thus, the treatment PF shows no positive employment effects, but it increases the benefit recipiency rate for women starting treatment later in their unemployment spell.

The evidence for SPST in figures 2.3–2.4 is much more positive and confirms the results in FS. After strong negative lock–in effects during a period of almost two years, we find positive and mostly significant medium– and long–run employment

[13]We discuss lock–in effects for the time it takes for the treated individuals to catch up with the nontreated individuals.

effects of around 10 percentage points (pp), which typically persist until the end of the observation period. The effects on BR are similar to PF, i.e. treatment increases BR in the short run, and the medium– and long–run effects are not significantly different from zero. The cumulated ER increases lie between 0 and 1.5 quarters. They are significant for stratum 1 and insignificant for the later strata. For men participating in SPST in stratum 1 we find the largest positive cumulated employment effects in this study. These men are on average 1.5 quarters longer employed during the course of six years since treatment start than comparable men who do not participate in a program in stratum 1. The average ER effects of SPST are highly significant and amount to about 10 pp in all cases. All cumulated BR effects are positive and significantly so for strata 2 and 3. The average BR effects are never significant. The effects for both genders are very similar.

For RT, the evidence in figures 2.5–2.6 is more mixed. As to be expected, we find the longest (typically lasting 10 quarters) and deepest lock–in effects for this treatment, with stratum 1 for men showing the strongest decline. The medium– and long–run ER effects are only significantly positive for males in stratum 1 and females in stratum 3. For women in stratum 1 the effects are sometimes significantly positive. The three other cases basically show insignificant ER effects in the medium– and long–run, although they are positive in most periods. Again, we find positive BR effects during the lock–in period and typically insignificant BR effects in the medium– and long–run for strata 2 and 3. For stratum 1 we see a medium– and long–run reduction, but which is only sometimes significant. Almost all of the cumulated ER effects are insignificantly negative, stratum 2 for men shows a significantly negative effect and stratum 3 for women an insignificantly positive one. Confirming the graphical evidence, the average ER effects are significant only for males in stratum 1 (around 12 pp) and females in stratum 3 (around 16 pp). All cumulated BR effects are significantly positive. The average BR effects are only significant for males in stratum 2 and 3.

No case in figures 2.1–2.6 shows significant differences in outcomes before the beginning of the unemployment spell. Since we include the employment history in the propensity score estimation, this is not a pre-program test of the CIA. But the results show that our matching approach balances well the employment history of treated and nontreated individuals. Note furthermore that lock–in effects last fairly long in comparison to results for West Germany, see Lechner et al. (2005a), LMW, FS, and Fitzenberger et al. (2006). A likely reason is that search frictions in the labor market are higher in East Germany compared to West Germany.

Overall, our results do not confirm the gender differences in the treatment effects as found in LMW. Neither for SPST, which comprises most of the long training as in LMW, and nor for RT, we find that employment effects are higher for females compared to males and that males show zero or negative long–run effects.[14] To explore reasons for the differences in results, we first would like to reexamine the evidence on gender differences in the content of training as reported in LMW, which the authors identify as a potential reason for the gender differences in the treatment effects. Programs are characterized by the target profession of training. This information is contained in table 2.5 stratified by gender, program, and stratum. Large differences show up between genders as also documented in LMW. PF for women mainly train in office professions (38%–48%) and in broader programs (20–27%), which can not be related to a specific profession. For female participants in SPST these fields are also the most important with 20–30% for office professions and 13–31% for broader programs. RT for women train mainly in service professions (17–28%), office professions (12–25%) and health professions (10–22%). For males, the programs PF and RT are dominated by target professions in *construction*, which have a share of at least 40%, and even 56% for men in RT in stratum 3. Metal professions are second most important for PF and RT in stratum 1 and 2 with about 25%. RT in stratum 3 trains only 12% in metal professions. SPST for men is concentrated in service professions (13–22%) and technical professions (13–19 %) for all strata. In strata 1 and 2 metal professions are most important with 27 and 23% and construction is also important with 13 and 17%. In the third stratum broad programs are most important with 32%. Thus, our data show similar gender differences in the content of training as reported by LMW.

Now, we explore further possible explanations of the differences in the estimated treatment effects for RT. We focus on RT because SPST differs from long training as defined in LMW and target professions in construction have a fairly small share in SPST. First, the differences to LMW are not due to the fact that LMW use a static evaluation approach, while we estimate the effects of treatment versus waiting. To investigate this, we reestimate the treatment effects in stratum 1 excluding the future participants in any training program from the control group (around 10% of the male

[14]As one exception, we find positive effects of RT for females and not for males in stratum 3. However, the number of treated males in stratum 3 is very small and the results in LMW correspond mainly to stratum 1 and 2 because the construction of the treatment sample in LMW oversamples early treatments, see discussion below.

and around 20% of the female controls are excluded, see additional appendix). The results for males basically do not change while the estimated treatment effects for females are reduced to some extent (these results are available upon request). Thus, the difference in evaluation approach should work in the opposite direction and can not explain the differences in the results. Second, since LMW suggest that males do not show positive long–run employment effects from RT because of the large share of target professions in construction, we estimate the treatment effects of RT for males separately with target profession in construction and in nonconstruction. We exclude the cases where the target profession is missing. The results (see additional appendix for details) clearly show that the employment effects for target profession construction are by no means smaller than for target profession nonconstruction. In fact, the point estimates for stratum 1 and 2 even suggest that in most cases the medium– and long–run employment effects are higher for target professions in construction (these differences are, however, not significant). Third, the differences in the sample construction (see table in additional appendix for a juxtaposition) between our paper and LMW show that LMW oversample early treatments. This should work in the opposite direction of the differences in the results, because in stratum 1 men but not women show positive employment effects for RT (see footnote 14). There are a number of further differences in the construction of the sample which, however, seem unlikely to explain the differences in results.

Concluding, we can not replicate the gender differences in results reported in LMW and we can not confirm differences in treatment effects by target profession as suggested by LMW. We have explored possible reasons to rationalize these differences but, unfortunately, the reason for these differences in results remains an open question.

2.6 Conclusions

Using a dynamic multiple treatment framework, this study analyzes the effects of three exclusive training programs for inflows into unemployment for the two years 1993/94. We evaluate medium– and long–run treatment effects both for employment and benefit recipiency up to 24–30 quarters after the beginning of the treatment depending on the starting date of the treatment and we distinguish by gender. Our results imply positive medium– and long–run employment effects for the largest program, Provision of Specific Professional Skills and Techniques (SPST), a program which involves sizeable off–the–job class room training. In contrast, practice firms show no positive employment effects and this holds also for retraining (the longest program) in four out of six cases. Furthermore, we do not find any of the three programs to reduce significantly the benefit recipiency rate in the medium and long run, in the short run all programs show the lock–in effect with an increase in the benefit recipiency rate, thus providing evidence for 'benefit churning' as in Kluve et al. (2004). The fact that we see increased ER and constant BR in the long run for SPST means that nonparticipation in the labor market went down. This suggests that such programs prevent its participants from leaving the labor force. Overall, the treatment effects are quite similar for females and males, thus, we can not confirm the gender differences found in Lechner et al. (2005b). Our evidence confirms the necessity to analyze long–term effects of sizeable training programs because all programs show strong negative lock–in effects in the short run. The positive assessment of SPST compared to practice firms is in contrast to the conventional wisdom in most of the literature. As two final caveats, an overall assessment of the microeconomic effects is not possible, because various necessary information for a comprehensive cost–benefit–analysis are lacking in our data set, and our paper ignores the likely general equilibrium effects of the large scale training programs implemented in East Germany.

2.7 References

Abadie, A. and G. Imbens (2006). "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica* 74, 235-267.

Abbring, J., and G.J. van den Berg (2003). "The Nonparametric Identification of Treatment Effects in Duration Models." *Econometrica* 71, 1491–1517.

Bender, S., A. Bergemann, B. Fitzenberger, M. Lechner, R. Miquel, S. Speckesser, and C. Wunsch (2005). "Über die Wirksamkeit von Fortbildungs- und Umschulungsmaßnahmen", *Beiträge zur Arbeitsmarkt- und Berufsforschung*, IAB, Nürnberg.

Bender, S., A. Haas, and C. Klose (2000). "IAB employment subsample 1975–1995", *Schmollers Jahrbuch (Journal of Applied Social Science Studies)* 120, 649–662.

Bergemann, A. B. Fitzenberger, and S. Speckesser (2004). "Evaluating the Dynamic Employment Effects of Training Programs in East Germany Using Conditional Difference–in–Differences." ZEW Discussion Paper No. 04-41, Mannheim.

BA ([Bundesanstalt für Arbeit] 1993, 1997, 2001). *Berufliche Weiterbildung.* Nürnberg: Bundesanstalt für Arbeit (various issues).

Bundesanstalt für Arbeit (2003), *Geschäftsbericht 2002*, Einundfünfzigster Geschäftsbericht der Bundesanstalt für Arbeit. Nürnberg: Bundesanstalt für Arbeit.

Fay, R. (1996). "Enhancing the Effectiveness of Active Labour Market Policies: Evidence from Programme Evaluations in OECD countries." Labour Market and Social Policy Occasional Papers, 18, OECD, Paris.

Fitzenberger, B., A. Osikominu, and R. Völter (2006). "Get Training or Wait? Long–Run Employment Effects of Training Programs for the Unemployed in West Germany." IZA Discussion Paper No. 2121.

Fitzenberger, B. and Prey, H. (2000). "Evaluating Public Sector Sponsored Training in East Germany.'" *Oxford Economic Papers* 52, 497-520.

Fitzenberger, B. and S. Speckesser (2007). "Employment Effects of the Provision of Specific Professional Skills and Techniques in Germany." *Empirical Economics* 32(2), 529-573.

Fredriksson, P. and P. Johansson (2003). "Program Evaluation and Random Program Starts." Institute for Labour Market Policy Evaluation (IFAU), Uppsala, Working Paper, 2003:1.

Fredriksson, P. and P. Johansson (2004). "Dynamic Treatment Assignment – The Consequences for Evaluations Using Observational Data." IZA Discussion Paper No. 1062.

Gerfin, M. and M. Lechner (2002). "Microeconometric Evaluation of the Active Labor Market Policy in Switzerland." *Economic Journal*, 112(482), 854–893.

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

Heckman, 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*, Vol. 3 A, Amsterdam: Elsevier Science, 1865–2097.

Horowitz, J. (2001). "The Bootstrap." In: J.J. Heckman, E. Leamer, editors, *Handbook of Econometrics*, Volume 5, Elsevier, Amsterdam, 3159-3228.

Hujer, R., S. Thomsen und C. Zeiss (2006). "The Effects of Vocational Training Programmes on the Duration of Unemployment in Eastern Germany" *Allgemeines Statistisches Archiv* 90(2), 299-321.

Imbens, G. (2000). "The Role of the Propensity Score in Estimating Dose–Response Functions" *Biometrika* 87, 706-710.

Kluve, J. H. Lehmann, and C. Schmidt (2004). "Disentangling Treatment effects of labor market histories: the role of employment histories." Discussion Paper, RWI, Essen.

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

Kraus, F., P. Puhani, and V. Steiner (1999). "Employment effects of publicly financed training programs - The East German Experience." *Jahrbücher für Nationalökonomie und Statistik*, 219, 216–248.

Lechner, M. (2000). "An evaluation of public sector sponsored continuous vocational training programs in East Germany." *Journal of Human Resources*, 35, 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.) (2000), *Econometric Evaluation of Active Labor Market Politics in Europe*, Heidelberg: Physica–Verlag.

Lechner, M., R. Miquel, and C. Wunsch (2005a). "Long–Run Effects of Public Sector Sponsored Training in West Germany." IZA Discussion Paper No. 1443.

Lechner, M., R. Miquel, and C. Wunsch (2005b). "The Curse and Blessing of Training the Unemployed in a Changing Economy: The Case of East Germany after Unification." Discussion Paper, University of St. Gallen.

Lubyova, M. and J.C. van Ours (1999). "Effects of Active Labour Market programs on the transition rate from unemployment into regular jobs in the Slovak republic." *Journal of Comparative Economics*, 27, 90–112.

Martin, J.P. and Grubb, D. (2001). "What works and for whom: A review of OECD countrie's experiences with active labour market policies." *Swedish Economic Policy Review*, 8, 9–56.

OECD (2005) "Labour Market Programmes and Activation Strategies: Evaluating the Impacts." Chapter 4 of Employment Outlook, OECD, Paris.

Puhani, P. (1999). "Evaluating active labour market policies - empirical evidence for Poland during transition." *ZEW Economic Studies*, 5, Physica, Heidelberg.

Rosenbaum, P.R. and D.B. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70, 41–55.

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

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

Schneider, H., K. Brenke, D. Hess, L. Kaiser, J. Steinwede und A. Uhlendorff (2006). "Evaluation der Maßnahmen zur Umsetzung der Vorschläge der Hartz-Kommission – Modul 1b: Förderung beruflicher Weiterbildung und Transferleistungen." IZA Research Report, No. 7, Bonn.

Sianesi, B. (2003). "Differential Effects of Swedish Active Labour Market Programs for Unemployed Adults in the 1990s." Discussion Paper, Institute for Fiscal Studies, London.

Sianesi, B. (2004). "An Evaluation of the Swedish System of Active Labor Market Programs in the 1990s." *Review of Economics and Statistics* 86, 133–155.

Smith, J.A. and P. Todd (2005). "Rejoinder." *Journal of Econometrics* 125, 365–375.

Speckesser, S. (2004). “Essays on Evaluation of Active Labour Market Policy.” Dissertation, University of Mannheim.

Wunsch, C. (2006). “Labour Market Policy in Germany: Institutions, Instruments and Reforms since Unification.” Discussion Paper, University of St. Gallen.

2.8 Appendix

2.8.1 Descriptive Statistics and Description of Data

Table 2.1: Participation in First Training Program for the Inflow Samples into Unemployment

Training Program	Frequency	Percent of inflow sample	Percent among treated
	Women		
Practice Firm	145	2.4	9.4
SPST	1,210	19.7	78.1
Retraining	195	3.2	12.6
No training program above	4,585	74.7	–
Total inflow sample	6,135	100	100
	Men		
Practice Firm	73	1.2	8.7
SPST	528	8.9	63.2
Retraining	234	4.0	28.0
No training program above	5,076	85.9	–
Total inflow sample	5,911	100	100

Remark: Programs that start before a new job is found are considered. We exclude training programs which start together with a job (like integration subsidies) or which involve a very small number of participants since they are not targeted on inflows into unemployment (as career advancement and German language courses).

Table 2.2: Number of Training Spells and Length of Unemployment before Program Start

	Practice Firm		SPST		Retraining	
	Women	Men	Women	Men	Women	Men
1–2 quarters	37	40	254	200	61 (61)	113 (107)
3–4 quarters	51	15	374	141	76 (75)	82 (79)
5–8 quarters	48	14	435	144	53 (53)	35 (33)
>8 quarters	9	4	147	43	5 (5)	4 (4)
Total	145	73	1,210	528	195 (194)	234 (223)

Remark: The time intervals indicate the quarter of program start relative to the beginning of the unemployment spell. The numbers in parenthesis for RT are participants who are less than 51 years old when entering unemployment.

Table 2.3: Elapsed Duration of Unemployment in Months at Beginning of Training Spell

	Practice Firm		SPST		Retraining	
	Women	Men	Women	Men	Women	Men
Average	10.9	8.1	12.8	10.4	8.9	6.8
25%–Quantile	5	2	6	4	4	3
Median	10	5	11	7.5	8	6
75%–Quantile	14	11	18	15	12	10

Table 2.4: Realized Duration of Training Spells in Months

	Practice Firm		SPST		Retraining	
	Women	Men	Women	Men	Women	Men
Average	6.5	6.1	9.1	8.8	18.7	17.3
25%–Quantile	6	4	6	4	15	12
Median	6	6	10	9	21	21
75%–Quantile	7	8	12	12	22	22

Remark: The duration of the training spell is defined as the number of months of continuous training. No interruptions are allowed. If in any month we do not identify the program we assume the program has ended the month before.

Table 2.5: Program fields of (target) profession

	Stratum	1	2	3	4	5	6	7	8	9	10	11	12	missing
		Program field (see below)												
Women														
PF	1	14	5	5	3	0	0	8	0	38	0	0	27	0
PF	2	6	4	6	2	4	2	4	4	43	0	6	20	0
PF	3	4	2	2	2	2	6	4	0	48	0	8	21	0
SPST	1	5	0	4	0	3	9	1	18	30	6	11	13	26
SPST	2	4	2	5	0	4	11	2	10	20	8	14	20	30
SPST	3	3	2	2	1	2	11	0	7	21	6	13	31	31
RT	1	6	0	4	2	0	10	2	8	12	20	28	8	18
RT	2	9	1	4	6	4	1	0	1	16	22	27	6	12
RT	3	12	2	4	8	4	8	0	0	25	10	17	8	9
Men														
PF	1	2	25	0	42	2	0	0	2	2	0	5	18	0
SPST	1	3	27	1	13	13	3	8	4	1	0	22	5	29
SPST	2	2	23	1	17	19	4	5	6	1	2	13	5	34
SPST	3	2	10	1	7	15	0	7	4	1	0	21	32	43
RT	1	3	27	0	43	2	1	7	0	0	2	12	3	17
RT	2	1	26	1	40	6	4	7	0	1	6	6	1	12
RT	3	0	12	0	56	3	0	3	0	6	3	15	3	3
Total		4	9	3	10	5	6	3	6	16	5	14	17	25

Remark: The table shows the distribution of the fields of profession for the programs by stratum and gender in percent of the nonmissing information. The fields are the following: 1 agriculture, basic materials, leather, textiles 2 metal 3 food 4 construction 5 technical 6 retail sales 7 transport 8 accounting 9 office 10 health 11 services 12 broader program. The last column gives the share of missing information.

2.8.2 Estimated Effects of Further Training Measures

Figures 2.1–2.6 display the estimated treatment effects $\hat{\theta}(k;\tau)$ on the horizontal axis against quarter $\tau \geq 0$ since the beginning of treatment or quarter $\tau < 0$ before the beginning of treatment. The time axis is divided into three parts by two vertical lines, which denote the last quarter before the unemployment spell starts and the treatment start $\tau = 0$, respectively. The left part shows the four quarters before unemployment starts, the middle part the gap between the beginning of the unemployment spell and the treatment start and the right part the time since treatment start.

Figure 2.1: Practice Firm (PF) for Women

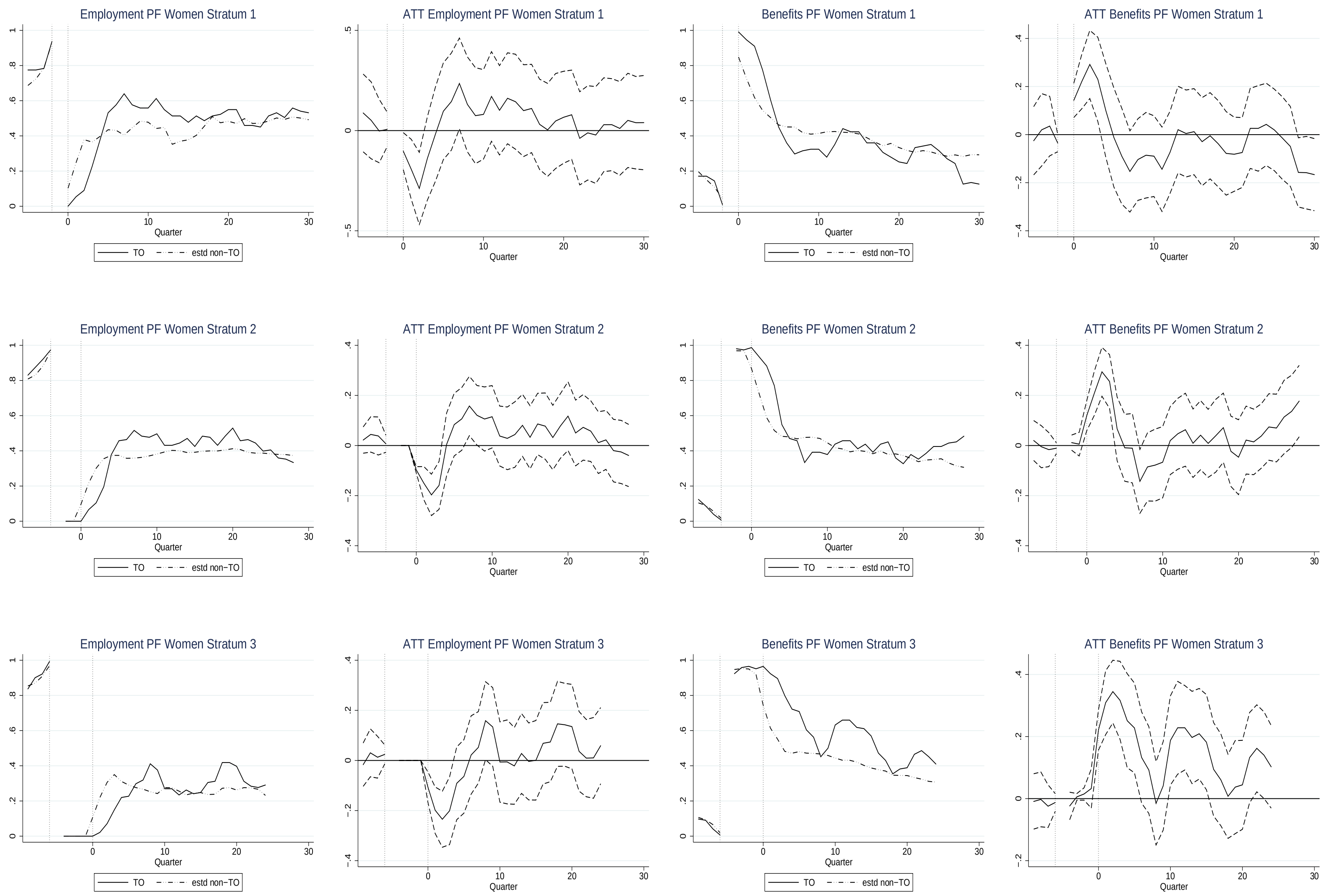


Figure 2.2: Practice Firm (PF) for Men

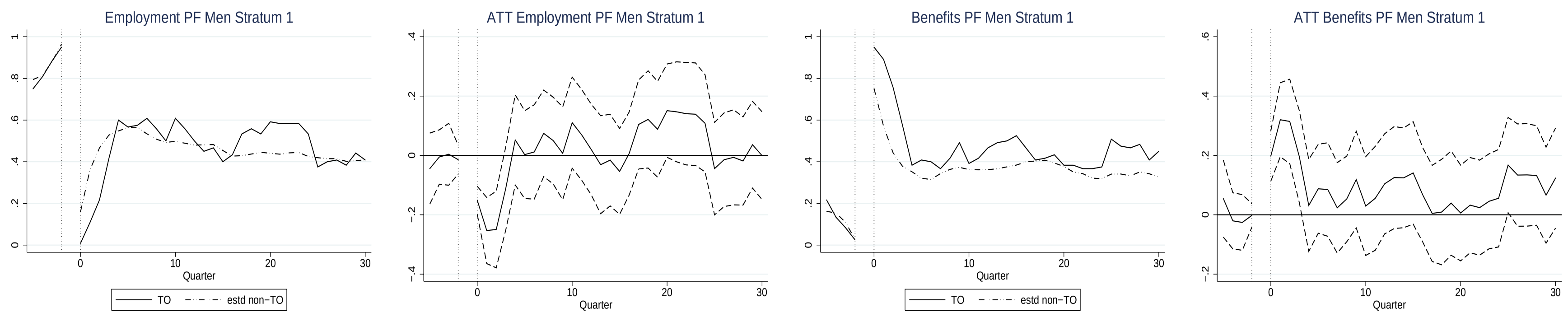


Figure 2.3: SPST for Women

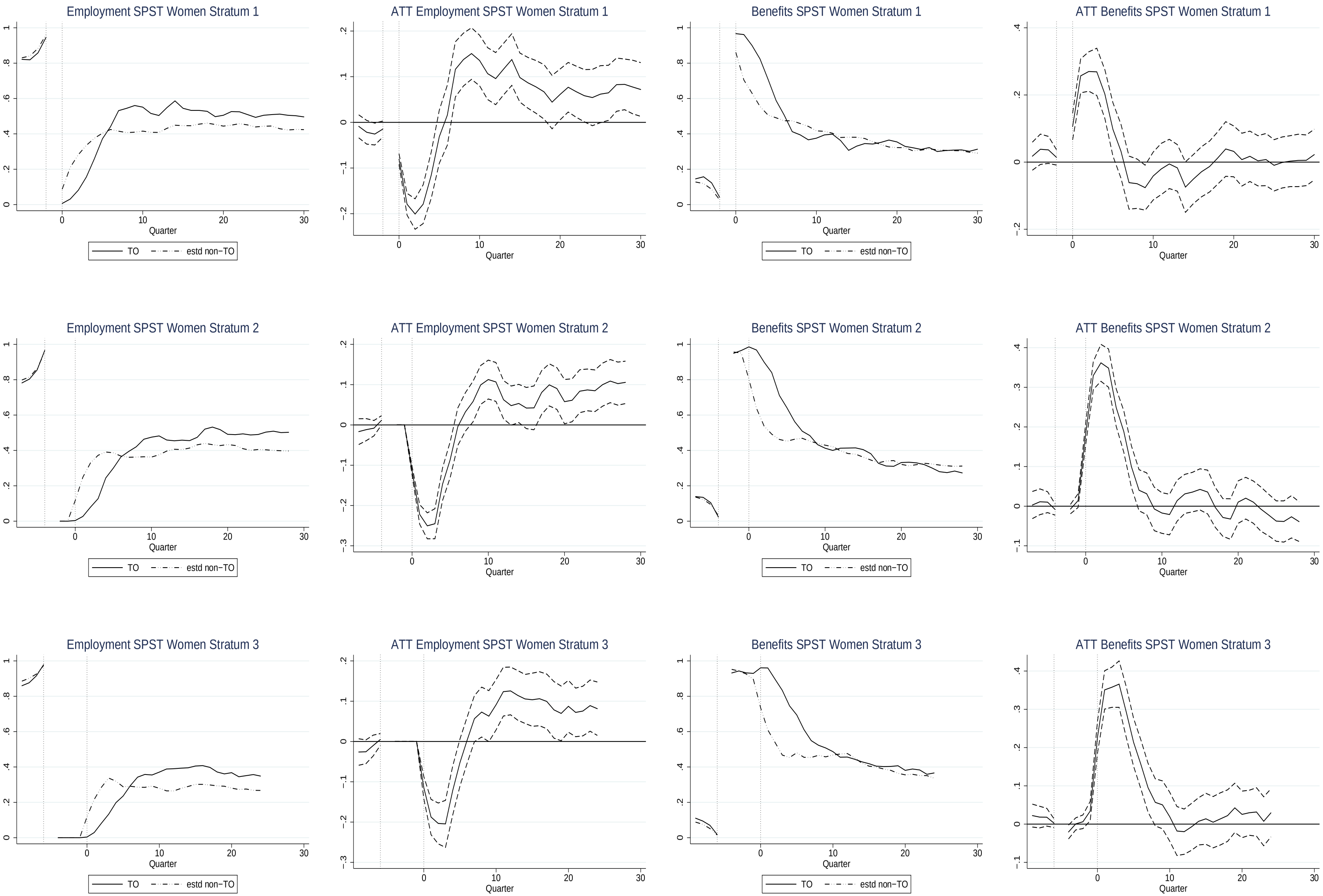


Figure 2.4: SPST for Men

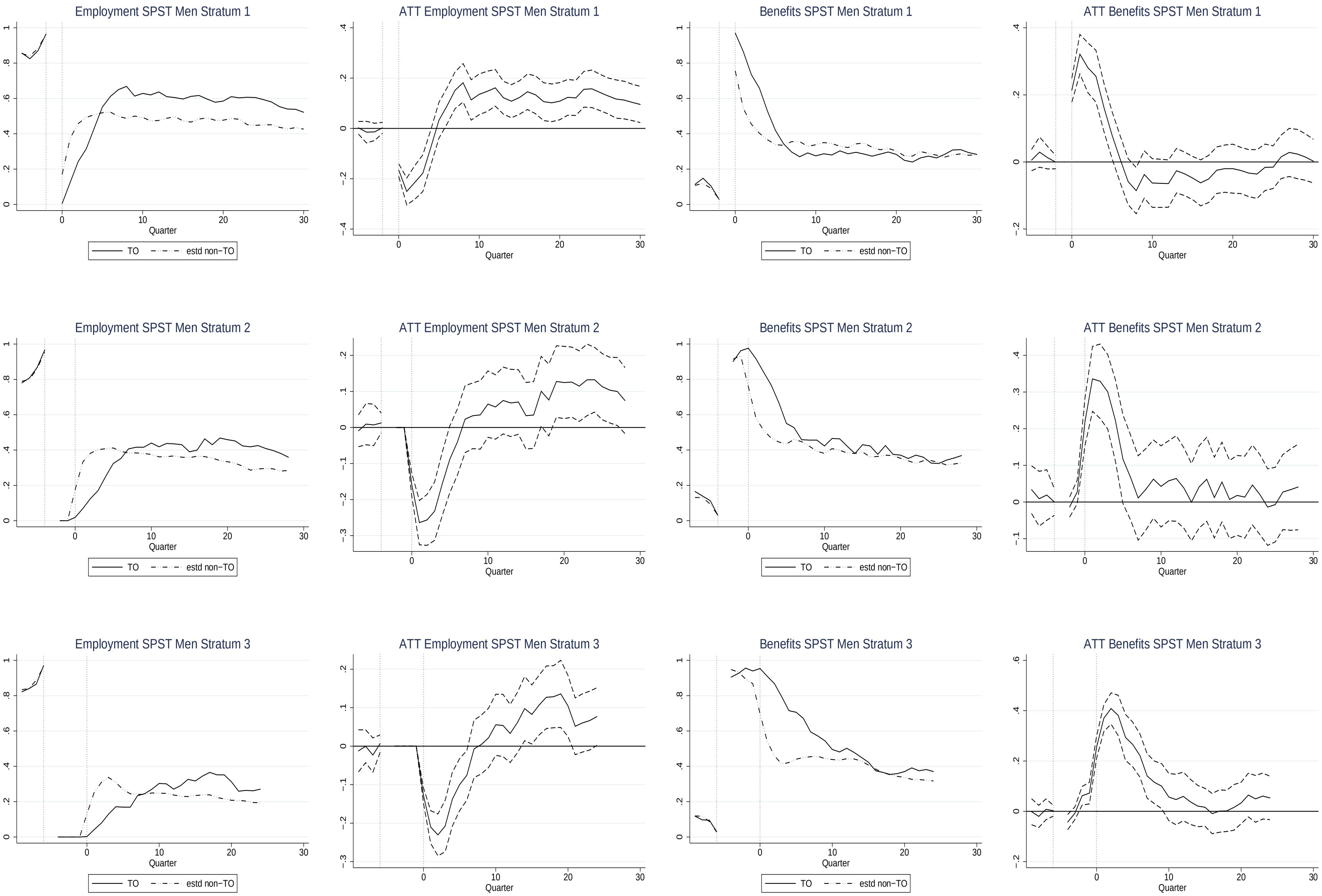


Figure 2.5: Retraining (RT) for Women

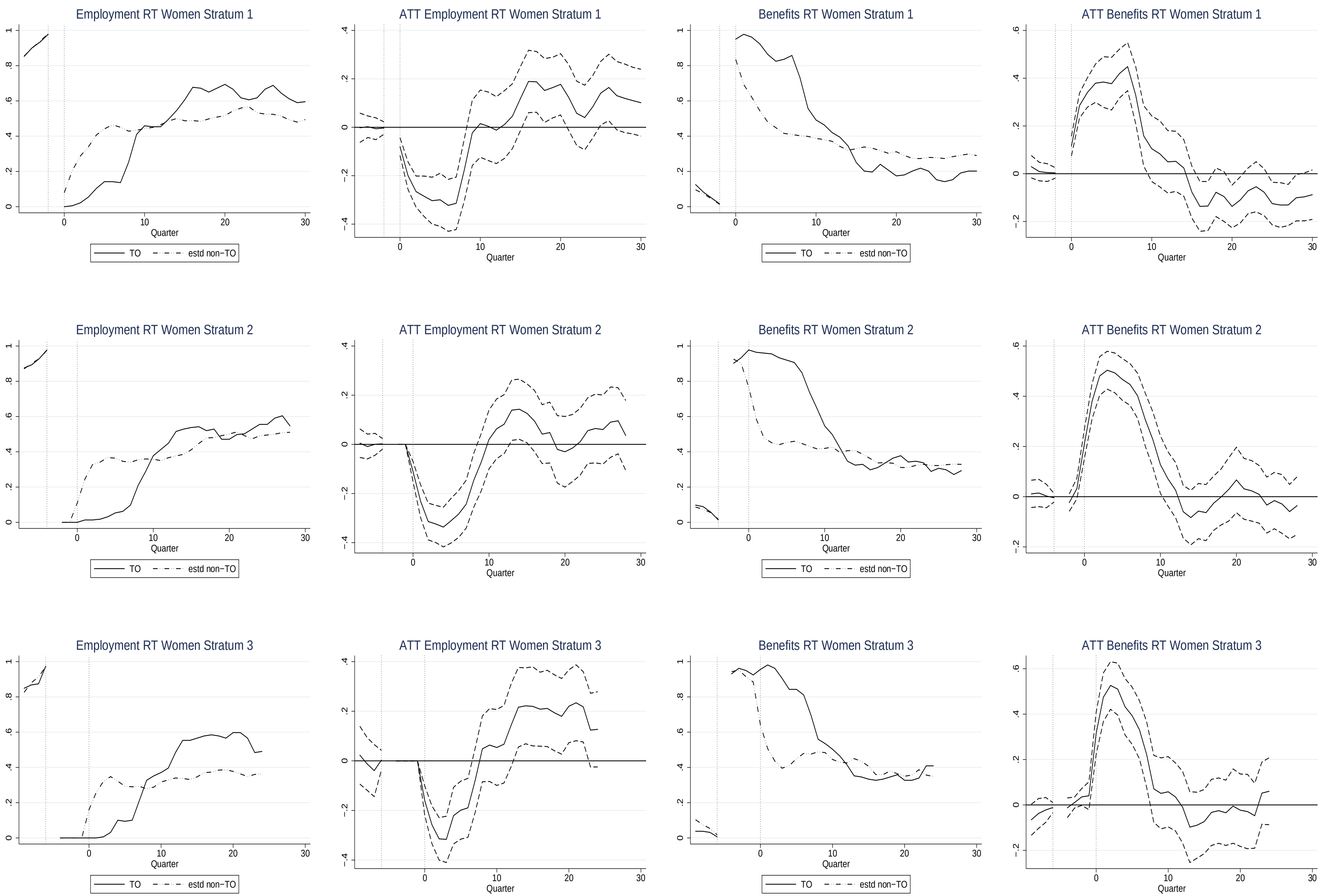


Figure 2.6: Retraining (RT) for Men

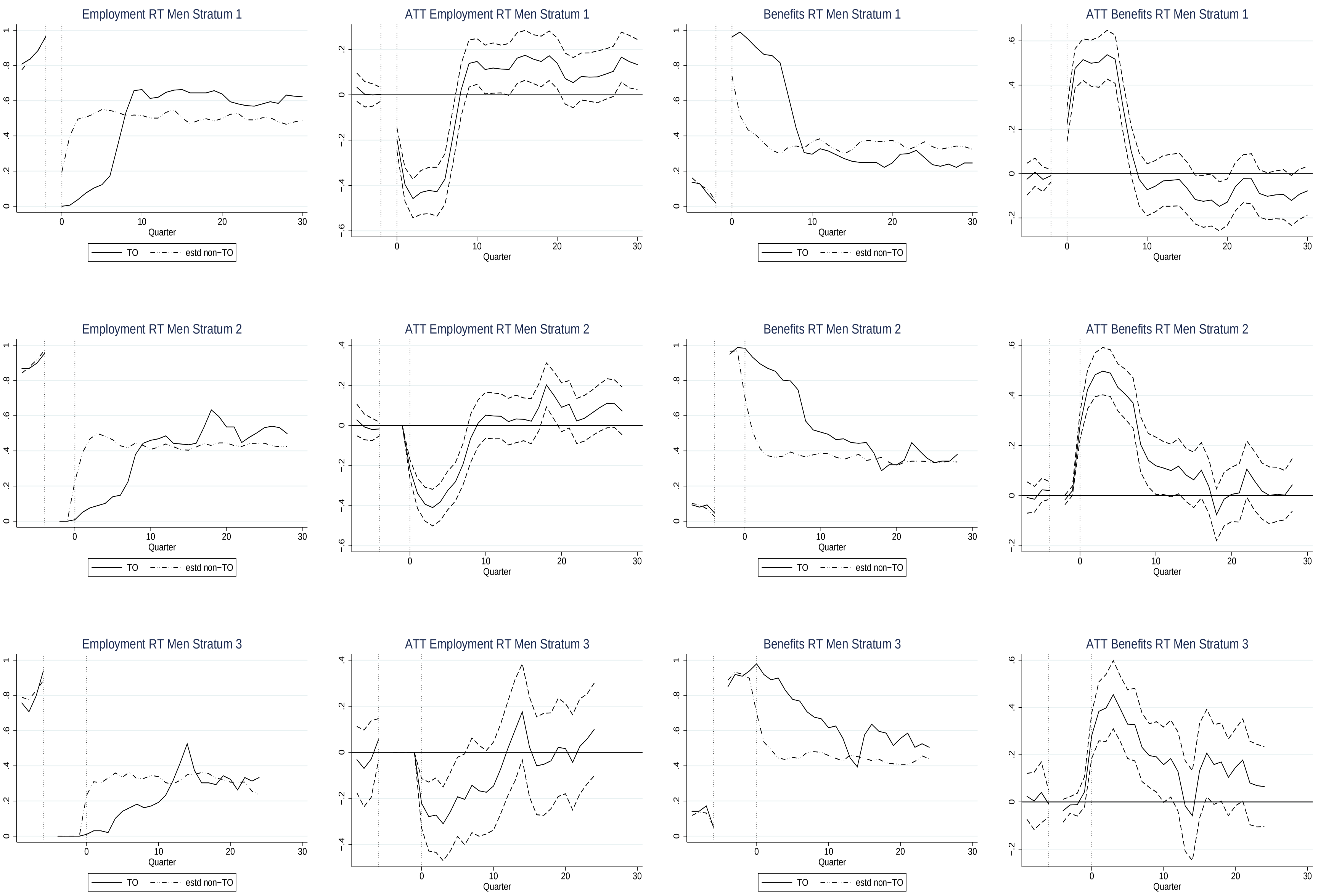


Table 2.6: Cumulated and Average Employment Effects

	Stratum	Cumulated Effects			Average Effects	
		Q0–Q7	Q0–Q15	Q0–Q23	Q4–Q23	Q8–Q23
Women						
PF	1	-0.27 (0.60)	0.70 (1.21)	0.98 (1.81)	0.085 (0.085)	0.078 (0.088)
PF	2	-0.25 (0.27)	0.31 (0.58)	0.88 (0.86)	0.074 (0.040)*	0.071 (0.044)
PF	3	-0.82 (0.37)**	-0.54 (0.77)	0.08 (1.19)	0.041 (0.056)	0.057 (0.061)
SPST	1	-0.65 (0.13)***	0.32 (0.26)	0.86 (0.38)**	0.075 (0.018)***	0.095 (0.020)***
SPST	2	-1.04 (0.10)***	-0.45 (0.22)**	0.15 (0.37)	0.049 (0.017)***	0.074 (0.019)***
SPST	3	-0.82 (0.17)***	-0.02 (0.33)	0.66 (0.51)	0.068 (0.024)***	0.092 (0.025)***
RT	1	-2.07 (0.25)***	-2.09 (0.54)***	-1.00 (0.84)	-0.008 (0.041)	0.067 (0.049)
RT	2	-2.15 (0.26)***	-1.80 (0.57)***	-1.61 (0.91)*	-0.032 (0.043)	0.034 (0.048)
RT	3	-1.74 (0.31)***	-0.70 (0.69)	0.89 (1.03)	0.097 (0.049)**	0.164 (0.055)***
Men						
PF	1	-0.62 (0.40)	-0.47 (0.71)	0.43 (1.05)	0.060 (0.048)	0.066 (0.052)
SPST	1	-0.61 (0.19)***	0.48 (0.39)	1.48 (0.56)***	0.114 (0.026)***	0.131 (0.027)***
SPST	2	-1.17 (0.24)***	-0.74 (0.49)	0.10 (0.74)	0.050 (0.035)	0.079 (0.036)**
SPST	3	-1.10 (0.18)***	-0.69 (0.38)*	0.09 (0.59)	0.044 (0.028)	0.074 (0.030)**
RT	1	-2.88 (0.30)***	-1.96 (0.56)***	-0.96 (0.82)	0.026 (0.037)	0.120 (0.040)***
RT	2	-2.54 (0.28)***	-2.36 (0.54)***	-1.64 (0.81)**	-0.014 (0.037)	0.056 (0.042)
RT	3	-1.88 (0.53)***	-2.13 (1.06)**	-2.20 (1.61)	-0.056 (0.073)	-0.020 (0.077)

Remark: The cumulated (average) effects are the sum (average) of the quarter specific average treatment effect on the treated over the respective quarters. $*$, $**$ and $***$ denote significance at the 10%, 5% and 1% level, respectively, and QiQj denotes quarter i to quarter j since beginning of treatment.

Table 2.7: Cumulated and Average Effects on Benefit Recipiency

	Stratum	Cumulated Effects			Average Effects	
		Q0–Q7	Q0–Q15	Q0–Q23	Q4–Q23	Q8–Q23
Women						
PF	1	0.73 (0.48)	0.27 (0.85)	0.02 (1.21)	-0.043 (0.057)	-0.045 (0.059)
PF	2	0.78 (0.30)***	0.73 (0.62)	0.86 (0.86)	-0.001 (0.040)	0.005 (0.045)
PF	3	1.89 (0.36)***	3.15 (0.70)***	3.83 (1.03)***	0.132 (0.048)***	0.121 (0.053)**
SPST	1	1.17 (0.20)***	0.82 (0.34)**	0.89 (0.51)*	-0.000 (0.024)	-0.017 (0.026)
SPST	2	1.80 (0.13)***	1.91 (0.24)***	1.91 (0.35)***	0.035 (0.016)**	0.007 (0.018)
SPST	3	2.06 (0.18)***	2.16 (0.33)***	2.34 (0.49)***	0.052 (0.023)**	0.018 (0.024)
RT	1	2.75 (0.26)***	3.47 (0.51)***	2.65 (0.71)***	0.077 (0.034)**	-0.006 (0.040)
RT	2	3.38 (0.25)***	3.94 (0.50)***	4.01 (0.76)***	0.122 (0.035)***	0.039 (0.040)
RT	3	3.21 (0.36)***	3.15 (0.74)***	3.01 (1.08)***	0.059 (0.050)	-0.013 (0.054)
Men						
PF	1	1.26 (0.39)***	2.01 (0.69)***	2.24 (1.03)**	0.060 (0.048)	0.061 (0.054)
SPST	1	1.26 (0.19)***	0.84 (0.35)**	0.57 (0.51)	-0.025 (0.023)	-0.043 (0.025)*
SPST	2	1.60 (0.30)***	1.94 (0.58)***	2.17 (0.84)**	0.050 (0.040)	0.036 (0.041)
SPST	3	2.34 (0.23)***	2.79 (0.46)***	3.01 (0.63)***	0.080 (0.029)***	0.042 (0.030)
RT	1	3.53 (0.34)***	3.32 (0.60)***	2.57 (0.81)***	0.044 (0.036)	-0.060 (0.038)
RT	2	3.38 (0.28)***	4.32 (0.48)***	4.55 (0.69)***	0.143 (0.032)***	0.073 (0.035)**
RT	3	2.80 (0.43)***	3.71 (0.71)***	4.82 (1.12)***	0.165 (0.051)***	0.127 (0.058)**

Remark: The cumulated (average) effects are the sum (average) of the quarter specific average treatment effect on the treated over the respective quarters. $*$, $**$ and $***$ denote significance at the 10%, 5% and 1% level, respectively, and QiQj denotes quarter i to quarter j since beginning of treatment.

Chapter 3

Regional Unemployment and Regional Mobility in West Germany

3.1 Introduction

The unemployment rates in West Germany exhibit strong regional variation. In 1997, the last year of this study, they range from a minimum of 4.8% in Erding to a maximum of 20.9% in Bremerhaven. These differences are quite persistent over time as for instance Büttner (1999) shows. Equalizing such regional disparities would be efficiency enhancing and regional mobility can be one factor in order to do so. Herzog et al. (1993) in their survey on migration as spatial job search consider regional mobility a macro-based measure of labor market efficiency. Blanchard and Katz (1992) show that in the US regional mobility is an effective mechanism in equalizing disparities in regional unemployment rates. In Europe this is less so the case, as Decressin and Fatás (1995) find.

In this paper I analyze the relationship between regional unemployment rates and regional mobility at the individual level in Germany. Are individuals in high unemployment areas regionally more mobile than individuals in low unemployment areas?

The literature on regional mobility is very extensive (for an overview on internal migration see Greenwood (1997)). So I will rather selectively discuss some papers relevant to this study, which analyze regional mobility at the individual level and consider regional unemployment rates as one factor. In an older survey of eight studies for the US, Herzog et al. (1993) conclude that higher regional unemployment induces higher migration, but admit the evidence to be not overwhelming. In a newer paper, Yankow (2002) focusses on the job search behavior of displaced workers in the US. He estimates a competing-risk hazard model for exits from unemployment into employment in the local labor market compared to exits into employment in other labor markets via migration. He finds that higher local unemployment rates are associated with a decreased likelihood of finding employment locally but are unrelated to the likelihood of finding employment via migration. So the proportion of exits via migration is higher.

For Germany I will briefly review four studies. Hatzius (1994) estimates yearly migration probabilities at the individual level in West Germany using GSOEP data. Migration is defined as migration between federal states. He finds regional unemployment to have an insignificantly (10%) positive effect on individual migration probability. Windzio (2004) analyzes regional mobility between Northern and Southern Germany using the IAB data for 1984 to 1997, the same data for the same period as this study does. In a hazard rate model he finds that higher regional unemployment at the district level lowers the migration probability, where migration is defined as mobility between northern and southern Germany. Arntz (2005) investigates the regional job search behavior of unemployed workers in Germany also using the IAB data. She uses a competing risk framework for exits into employment within the same region as opposed to exits into employment into other regions (comparable to Yankow (2002)). Her measure of local economic conditions is not purely the unemployment rate, but the unemployment vacancy ratio. She finds that a higher unemployment vacancy ratio increases the migration hazard and lowers the rate of finding jobs locally. Lastly, Schündeln (2007) looks at migration within Germany using microcensus data. He estimates conditional logit models of the migration decision between German federal states. He finds migration to be responsive to the state unemployment rate. A higher state unemployment rate lowers the probability that individuals stay within the state or move towards the state. Summarizing the evidence for Germany, three studies find a moderately positive relationship between regional unemployment and regional mobility, but Windzio (2004) finds the contrary. This conflicting evidence motivates further investigation.

In the previous literature the regional unemployment rate is treated as one factor among others influencing regional mobility. The contribution of this study is its sole focus on the link between the regional unemployment rate and regional mobility. My specific approach in this paper is to analyze regional mobility based on job changes. The measure for regional mobility is the share of job changers which starts a new job in another region. The first question investigated is if regional mobility is higher in regions with higher unemployment. Indeed I find that regional mobility is higher in regions with a higher regional unemployment rate. A higher share of job changers coming from high unemployment regions also leaves the region. A second question analyzed is, if the destination choice of those who are regionally mobile is lead by differentials in unemployment rates. Only if those who are regionally mobile move towards regions with lower unemployment rates, regional mobility can help in equalizing unemployment rate differentials. I find that on average the regionally mobile move towards regions with lower unemployment rates.

I study job changes of men in West Germany between 1984 and 1997 using IAB employment register data. My intention is to explore the relationship between regional unemployment and regional mobility in a simple and descriptive fashion. It is beyond the scope of this study to model migration decisions. This would also require information unavailable in the IAB data about the household context and home ownership.

The remainder of this paper is organized as follows. The next section describes the data and the constructed sample. Section 3.3 gives the results and section 3.4 concludes.

3.2 Data and Sample

3.2.1 Data

Basic data set for this study is the IABS employment subsample (*IAB Beschäftigten-stichprobe, IABS*) of the Institute for Employment Research (IAB) for the time period 1975 to 1997, for details see Bender et al. (2000). The IABS is a 1% random sample of employment subject to social insurance contributions and covers 591,627 individuals. For those individuals in the sample the IABS comprises all employment spells between 1975 and 1997 as well as all spells during which the individuals have received unemployment benefits. In effect the IABS contains complete employment histories with the exception of self employment periods since the self employed do not have to pay social insurance contributions. For the same reason lifetime civil servants are not included in the IABS.

The regional information in the IABS is the district (*Kreis*) where the employer is located. Western Germany consists of 327 districts. So, the districts give quite a detailed picture about work place locations. However, the data include no information about the employees residential locations. As a consequence, regional mobility in this paper always means geographical mobility between workplaces and not residential mobility. I can not distinguish whether individuals actually migrate or commute. Since the focus of this study is the mobility of labor, this is only a small drawback. The question is if labor is geographically mobile and for this study it is of minor importance if this mobility is achieved by migrating or by commuting. As a measure of distance between two districts I calculate the straight-line distance between the two district capitals by using latitudes and longitudes.

The unemployment rate at the district level is provided by the Federal Employment Agency (*Bundesagentur für Arbeit, BA*). I use yearly averages of the district unemployment rates, because the rates are given by the BA in varying frequency. Since this study is limited to males it would be desirable to have unemployment rates for men. Unfortunately, for the time period under consideration unemployment rates at the district level differentiated by sex are not available. So I have to use unemployment rates for men and women together as a proxy.

3.2.2 Sample

The observational units to be analyzed in this study are job changes. The question is if the job changes take place within regions or between regions. Using the IAB data I have constructed a sample of job changes. Since the IABS is a representative sample of all individuals who have been employed subject to social insurance contributions at least once between 1975 and 1997, the job change sample is representative for all job changes of those individuals. Note, that as a consequence the sample is not representative for individuals. If some individuals change jobs more often than others do (for instance the younger) than these individuals contribute relatively more observations to the job change sample. A conventional alternative to the job based analysis of regional mobility would have been to keep the individual as the observational unit and use the IABS as a panel. Limitations of the IAB data prevented me from doing so. The geographical information (the employers location) is only available for periods of employment, but not for periods of non- or unemployment. Aggregating the spell based IABS into a yearly (or monthly) panel in order to estimate person based migration rates would thus have resulted in missing values for periods of un- or nonemployment. The job based analysis chosen, however, has the advantage that geographical information is always available for the jobs. Another alternative would have been to use a different data set, the GSOEP, the main household panel for Germany. But since regional mobility is a rather rare event I consider the much larger sample size of the IABS an important advantage.

Now I come to the details of the sample of job changes constructed from the IABS. This job change sample comprises job changes of men aged 20 to 65 during the years 1984 to 1997 within West Germany. The first year, 1984, is the first year for which district unemployment rates are available and the last year, 1997, is the last year covered by the IABS version used for this study. I exclude Berlin and East Germany in order to study a homogeneous and connected area. The job change sample contains direct job-to-job transitions, in which the new job starts within 7 days after the old job ends, as well es job changes characterized by an un- or nonemployment period of up to two years between both jobs. I exclude job changes with an intervening period of more than two years, because within such a long time it is more likely that some relevant event occurred, which is not recorded in the data, like for instance self-employment. The IAB data record many employment spells which last only for a few days. A typical example are engagements of artists. In order to obtain a sample of job changes between relatively stable jobs I restrict the sample to job changes between jobs which each last longer than 31 days.

Next I will explain how I define job changes. There are two classes of job changes in the sample. The first class are plant changes. This is determined by the IAB data, which identify firms at the plant level. The second class are recalls. I view recalls as job changes which result in a new job at the previous firm, but could have resulted in a new job at another firm. Since job changes of those, who expected to be recalled by their old firm but ended up at another firm, are by construction included in the first class of job changes, I decided to include recalls in the job change sample as well, since I consider expected recalls and realized recalls as quite comparable. Recalls make up 17% of the total sample.[1]

An overview about the constructed job change sample can be found in table 3.1. The sample contains 308,841 job changes from 115,548 men. Looking at the geographical distance between both jobs, about 62% of the job changes take place within the same district denoted by "0 km".[2] Distances between 1 and 50 km are covered by 22% of the job changes and distances between 51 and 100 km by 6%. Job changes with distances of more than 100 and up to 250 km make up 5% of the sample and 4% of all job changes are in the category with more than 250 km. In order to simplify the following analysis I aggregate the five distance categories into just two, referred to as changes within regions and changes between regions. All job changes with distances of up to 100 km are called changes within regions and changes with distances of more than 101 km are called changes between regions. Using this binary classification there are 90.7% changes within regions and 9.3% changes between regions.

Differentiating the job changes by potential events between the old and the new

[1]Note that recalls necessarily are job changes within a region, because the new employer is the old one. Excluding recalls thus would increase measures of regional mobility.

[2]This category also contains changes between adjacent districts as long as they share the district capital.

job results in three types of job changes. The first and largest type with 44% are direct changes without an interruption of more than 7 days between the old and the new job. The other two types are characterized by an interruption of more than seven days between both jobs. The second type of job changes is described by an intervening period of unemployment with receipt of unemployment benefits for at least one day and comprises 35% of the sample. The third type with the remaining 21% is characterized by a period of non-employment without reception of unemployment benefits. The composition of this third type is potentially quite diverse, which makes the interpretation difficult. Individuals can for instance by unemployed without benefit payments or be self employed, which both is not recorded in the data.

An overview about the characteristics of the job changers is given in table 3.2. They are rather young, 45% in the sample are between 20 and 29 years old at the time of the job change.[3] The dominant education category is a vocational training degree with 77%. No vocational training degree have 16%, a university or technical college degree have 8%. Since the education variable given in the IABS has quality deficiencies, I applied the imputation procedure (IP2B) described in Fitzenberger et al. (2006) to improve its quality. In the sample, married are 41% and foreign 12%. Migration and employment history is captured by three variables. In the last two years before the job change considered, 7% have been regionally mobile (had at least one job change which also was a regional change), 13% have been recalled by an old employer and 29% have been unemployed with payment of benefits.

Several variables describe the old job and the old employer: employment status, earnings, tenure and firm size. Occupations and industries are accounted for with two sets of dummy variables. Occupations, which are given in the IAB data at the two-digit level, were aggregated into 25 occupations (table 3.3). The classification is adopted from Fitzenberger and Spitz (2004). Main occupations in the sample are transportation occupations with 14 percent of the observations and mechanics as well as construction related occupations each with 12 percent of the sample. The two-digit level industries were grouped into 18 industries (table 3.4). The largest

[3]Note, that the time of the job change has to be defined, in case the ending date of the old job is not the starting date of the new job. In my sample the intervening period can be up to two years. I define the time of the job change as the time when the new job starts. My reason for doing so is that I want it to be the date of the actual migration in case the job change is accompanied by migration. The time of the actual migration is also unobserved, but is likely to be (close to) the time when the new job starts. Gregg et al. (2004) argue for the UK that it is very rare for the unemployed to move before they have found a new job.

industry in the sample is construction with 18 percent of the observations[4], followed by wholesale and retail trade with 13 percent.

3.3 Results

3.3.1 Regional Mobility and Regional Unemployment

The first question to be analyzed is how the regional unemployment rate influences regional mobility. Does a higher regional unemployment rate increase regional mobility? I operationalize this question by looking at the probability that a job change is also a regional change, i.e. the geographical distance between both jobs is larger than 100 km. I analyze the share of regional changes among job changes in a linear probability model, depending on the regional unemployment rate and other covariates. The coefficient of interest is the coefficient of the district level unemployment rate in the district where the old job is located measured in the year of the job change. A positive coefficient would show that the probability of outmigration is higher in regions with higher unemployment. In the estimations I include control variables for (i) personal characteristics, (ii) attributes of the last job, (iii) the employment and migration history and (iv) regional information as well as year dummies. The consideration of the regional situation is done in two ways. In the first alternative I limit myself to the use of just one urban dummy for job changes originating in cities of more than 100,000 inhabitants. In the second alternative I include a full set of dummies for all 327 districts of the old job. This is important if there are permanent differences between districts in regional mobility (for instance due to differential road and rail links or central opposed to peripheral location) and these differences are correlated with differences in unemployment rates. Then district dummies would take up the permanent variation in the unemployment rates between districts and the coefficient on unemployment would be driven by variation in unemployment within districts over time. It would measure the relationship between regional mobility and relative regional unemployment, where relative means relative to the district specific average unemployment rate over time.

As a starting point I analyze the relationship between the unemployment rate and

[4]Construction as an industry has a notably larger share than construction related occupations since the occupation classification is finer than the industry classification and some occupations like painters are not contained in the general group construction related occupations.

regional mobility in a set of regressions without or with a very few covariates. Table 3.5 contains the estimates. The dependent variable is a dummy for regional mobility multiplied by 100, so that the estimated coefficients can be interpreted as percentage points of mobility. Regressing mobility on the unemployment rate at the district level (*alq*) gives an insignificant[5] coefficient (column 1). Hence overall, there is no significant correlation between the regional unemployment rate and regional mobility. One can see the regional unemployment rate as the sum of the national unemployment rate and the deviation between the regional unemployment rate and the national one. This difference is a relative regional unemployment rate, relative to the national one. It is quite likely that the national and the relative regional unemployment rate relate differently to mobility. Regressing mobility only on the national unemployment rate (*alqwd*) results in a significantly negative coefficient of -0.14 (column 2). If the national unemployment rate is 1 percentage point higher, mobility is 0.14 percentage points lower. Regional mobility shows to be procyclical, as commonly found in the literature, see Pissarides and Wadsworth (1989) or Greenwood (1997). The regression of mobility on both, the regional and the national unemployment rate, shows that both coefficients go in different directions (column 3). The regional unemployment rate has a significantly positive coefficient with 0.046 and the national unemployment rate a significantly negative coefficient with -0.192. This shows that the national unemployment rate is negatively correlated with mobility but relative regional unemployment is positively correlated. A slightly more general specification, in which year dummies instead of the national unemployment rate take up business cycle effects (column 4) results in the same coefficient for the regional unemployment rate with 0.046. Later on I will generally include time dummies instead of the national unemployment rate. Nevertheless, the interpretation of the coefficient for the district unemployment rate remains. It describes relative regional unemployment, relative to the national unemployment rate.

West Germany is quite heterogeneous and region specific effects might be correlated with regional mobility as well as with regional unemployment. In order to abstract from such district specific effects I include a complete set of district dummies in a fifth regression as described above (column 5). In this case the interpretation of the coefficient on the regional unemployment rate is relative in a double sense.

[5]The criterion for significance is significance at the 5% level. Reported coefficients without indication of significance are significant unless otherwise stated. The regression standard errors are robust standard errors clustered at the person level.

First it is relative to the national unemployment rate because time dummies are included and second it is relative to the district specific average deviation from the national unemployment rate because district dummies are included. The estimated coefficient is 0.154, which is three times larger than the coefficient without district specific effects. This shows that within districts higher unemployment is correlated with higher mobility, but between districts it is negatively correlated.

Regional Mobility and Regional Unemployment

Now I come to the main results from the estimations with covariates. They can be found in table 3.6. The left column shows the estimates without district dummies and the right column with district dummies. It can be seen that a one percentage point increase in the regional unemployment rate is related to an increase in regional mobility by 0.09 percentage points (pp) in the model without district dummies. This means that regional mobility is higher in districts with higher relative unemployment (relative compared to the national average since year dummies are included). In the right column we see that additionally controlling for district specific effects results in a coefficient of 0.199. Compared to the coefficient without district dummies this is twice the size, but the absolute size is still small. Even a 5 percentage point higher regional unemployment rate would mean a just one percentage point higher regional mobility. This has to be compared to an average rate of mobility of 9.3 percent.

Regional Mobility and Other Factors

Table 3.6 also shows how regional mobility varies with the control variables. For the sake of brevity I will limit the discussion to some important covariates and only report the coefficients for the estimation with district dummies. The coefficients for the estimation without district dummies are of comparable size. The by far largest coefficient has the dummy for regional mobility in the past, which can be seen as a lagged endogenous variable. Those men, who have had a regional change in the last two years before the job change considered (*mexp100y2*), have a 33.0 pp higher probability that the actual job change is also a regional change than those who have not been mobile in the last two years. Possible reasons include return migration (Dustmann (2003)) and generally higher levels of regional mobility for those who have been geographically mobile before. Men who have been unemployed in the last two years with receipt of unemployment benefits (*LEDexpy2*) are 2.9 pp less likely to be regionally mobile. Education is the second most important factor. Those with a university degree (*edu_h*) are 7.5 pp more likely to be regionally mobile than those with a vocational training degree (*edu_m*), whereas the difference between those with and without a vocational training degree is only 0.6 pp. Also important is the occupational status in the last job. White collar employees (*stib4*) are 3.7 pp more mobile than skilled blue collar employees. Tenure in the old job (and hence in the region) is negatively correlated with regional mobility. One year more tenure (*tenure*) means 0.3 pp less mobility. The relationship between age and mobility as measured in this framework is astonishingly small. Compared to men in their twenties men in their thirties (*age3039*) are 0.9 pp more likely to change the region when they change their job and men aged 50 and above (*age5065*) are 0.6 pp less likely. Married men are around one pp less mobile than unmarried men. Foreigners are 0.4 pp more mobile than natives. This is a comparatively small difference in contrast to Schündeln (2007), who finds (recent) immigrants to be considerably more mobile.

It is also informative to compare the order of magnitude of the coefficients for the covariates just discussed with the coefficient for regional unemployment. It shows that other factors than regional unemployment are considerably more important in explaining regional mobility at the individual level, notably previous regional mobility, education and employment status.

Regional Mobility and Regional Unemployment within Subgroups

In this subsection I want to explore if and how the relationship between regional mobility and regional unemployment differs between subgroups of job changes or job changers. I do so by introducing interaction effects in the analysis as well as by conducting the analysis separately for certain subgroups.

Results for estimations with interaction effects, which otherwise use the same specification as in table 3.6, are contained in table 3.7. Regarding age, the interaction of the regional unemployment rate with dummies for four age groups shows, that the relationship between regional unemployment and regional mobility is only significant for those below the age of 40. For men aged 40 and above no significant relationship is found. Accordingly are the estimated coefficients for regional unemployment for those below age 40 up to 50 percent higher than in estimations where regional unemployment is not interacted with age. A likely interpretation for the difference by age would be that regional unemployment works as a push factor out of a region for the young but not for the old because the young have a longer time horizon. They can gain from better economic conditions elsewhere for a longer period. Above we have seen that previous mobility is the single most important predictor for regional mobility. However, the relationship between regional mobility and regional unemployment does not differ by whether or not someone has been mobile before. The interaction effect for regional unemployment and the dummy for mobility within the last two years is insignificant. In the last subsection we have also seen that tenure is negatively related to regional mobility. Interacting the regional unemployment rate with a dummy for at least one year of tenure results in a significantly positive coefficient. So the relationship is stronger for those with more tenure. The time period under consideration, 1984–1997, can be divided into the years before and after German reunification. The reunification is important for this study of mobility within West Germany because it was accompanied by a large migration wave from East to West Germany and because of its macroeconomic implications in general. Interacting the regional unemployment rate with a dummy for the years until 1989 and another one for the years from 1990 onwards we see the significantly positive relationship with regional mobility only for the eighties, in the nineties no significant relationship can be found.[6]

Now I want to examine three subgroups of job changes in some more detail: direct job changes, job changes with an intervening period of unemployment and job changes with an intervening period of nonemployment. The incidence of mobility is different between these three subgroups. Among the direct job changers 11.5% also change the region, whereas among those who experience unemployment between the jobs only 6.5% change the region. The probability for those with a spell of nonemployment is in between with 9.8%. I am interested in how the relationship between regional unemployment and regional mobility differs between the three subgroups. Therefor I repeat the main analysis from above for each subgroup. Table 3.8 contains the results. Looking first at direct changes, it can be seen that the estimated coefficients for the district unemployment rate are somewhat higher than in the full sample. In the estimation without district dummies the coefficient is 0.099 compared to 0.090

[6]Results for interaction effects with other covariates (including education, employment status and nationality) can not be discussed in detail in order to keep the paper short. Regarding education and employment status I find no differences in the relationship between regional unemployment and regional mobility. Regarding nationality, foreigners are significantly more responsive to unemployment rate differentials than Germans.

and in the estimation with district dummies the coefficient is 0.267 compared to 0.199. In the second group of those, who experience personal unemployment, we have a different picture. The coefficient for the district unemployment rate in the estimation without district dummies is insignificant. In the estimation with district dummies the estimated coefficient is 0.161. This is somewhat smaller than in the full sample. The last group are those, who have a spell of nonemployment. The estimates for the regional unemployment rate are 0.234 for the estimation without district dummies and 0.404 with district dummies. This is each about twice the size of the coefficients in the full sample. Summarizing the evidence, we have the weakest relationship between regional unemployment and regional mobility for those, who are personally unemployed. The strongest relationship can be found for those who have a spell of nonemployment and the direct job changers are in between.

Finally, as a robustness check, I have analyzed the sensitivity of the relationship between regional unemployment and regional mobility to the definition of regional mobility. Instead of defining regional mobility as job changes which cover geographical distances of more than 100 km I also tried 50 km, 75 km, 125 km and 150 km as alternative cut-off points. The two cut-offs 75 km and 125 km, which are closest to the one used, both give quite comparable results to the ones presented. With some minor exceptions this is also true for the cut-off 150 km. Taking 50 km as the cutoff does not in general repeat the results from the cutoff 100 km. I take this as evidence that geographical distances of slightly more than 50 km are more likely to be covered within regions than between regions.

3.3.2 Direction Index

In this second part of the analysis I want to look at the destination choice of those who are regionally mobile. This means that now I concentrate on the 9.3% of the job changes which are regional changes. Table 3.9 gives an overview. The question I want to explore is if the movers move into the direction of lower unemployment. In order to do so, I take the regional changes i and compare the regional unemployment rate in the destination district $udestination_i$ with the regional unemployment rate in the origin district $uorigin_i$. I call this difference $directionindex_i$:

$$directionindex_i = udestination_i - uorigin_i.$$

The sign of the index indicates the direction of a regional change. If it goes into the direction of lower unemployment the index is negative, if it goes into the direction of higher unemployment the index is positive. Table 3.10 shows the average direction index. In the eighties, between 1984 and 1989, the average direction index is -0.36. This means that for all regional changes the unemployment rate in the destination districts is on average a third of a percentage point lower than in the origin districts. So we see a significant relationship between the destination choice of the regionally mobile and the regional unemployment rate and it goes in the right direction in order to equilibrate regional disparities in the unemployment rate. In the nineties, from 1990 to 1997, the average direction index is -0.08. It is much smaller than in the eighties. The table also shows the direction index for the three subgroups of job changes. The average direction index for direct job changes is -0.30 in the eighties and -0.07 in the nineties. This is only slightly different from the full sample. For job changes with an intervening period of unemployment we have an index of -0.61 in the eighties and -0.16 in the nineties. This is each about twice the size than for the full sample. And for the last group, job changes with a period of nonemployment we have a value of -0.19 in the eighties, which is roughly half the index in the full sample, and an insignificant index for the nineties. Taken together, we see in all three groups a much stronger relationship between the direction of regional changes and differences in regional unemployment rates in the eighties than in the nineties. Between the three groups, the unemployed show the strongest tendency to move towards regions with lower regional unemployment rates. Job changers with nonemployment have the weakest tendency and direct job changers lie in between.

It remains to be discussed how the direction index differs with sociodemographic covariates. In order to investigate this question I have regressed the direction index on sociodemographic covariates and dummies for the eighties and nineties separately for direct changes, changes with an unemployment period and changes with a nonemployment period. The results are shown in table 3.11. It can be seen that for all three groups of job changes age, education, marital status and nationality have no explanatory content for the direction index. However, the migration experience dummy has in all three groups a positive coefficient in the order of magnitude of the negative dummy for the eighties. This means that the overall tendency of regional changes to go towards regions with lower unemployment is not present for those who have recent migration experience. One likely explanation for this fact is return migration. Since return migration goes into the opposite direction than the original migration every association between the direction of migration and regional unemployment is mitigated by return migration. The relationship between the direction index and some remaining covariates shows notable differences between the types of job changes. Being in an apprenticeship or having just finished one (*stib0*) has a sizable negative coefficient for job changes with unemployment with -0.96 and to a lesser degree also for job changes with nonemployment with -0.59 but is insignificant for direct job changes. This shows, that those who change the region at the beginning of their career and in combination with a career interruption through unemployment or nonemployment have a stronger tendency to move towards regions with lower unemployment than job changers later in their career. This differentiation between apprentices and men later in their career can not be found for direct job changes. It is likely that these are more led by individual opportunities. Finally there are two factors which are only significant for job changes with unemployment. First, unskilled blue collar employees (stib1) have a coefficient of 0.51, which lessens their average tendency to move towards regions with lower unemployment compared to skilled blue collar employees. And second, an unemployment benefit payment in the last to years (*LEDexpy2*) has a negative coefficient of -0.43, which means that the drift towards regions with lower unemployment is stronger for unemployed with a history of unemployment.

Again, as a robustness test, I have also checked the sensitivity of the direction index to the definition of regional changes. I have repeated the analysis for all job changes which cover distances of at least 50, 75, 125 and 150 km, respectively. The results for the cut-offs 75, 125 and 150 km are quite comparable to the results shown for the cut-off 100 km. This is mostly also true for the cut-off 50 km. In general, the magnitude of the index is the larger the higher the minimum distance for regional changes is. This means the tendency to move towards regions with lower unemployment rates is somewhat stronger for job changes which cover longer distances than for job changes which cover shorter distances.

3.4 Conclusion

Motivation for this study are large and persistent regional differences in unemployment rates in West Germany. Equalizing this differences would be efficiency enhancing. Here regional mobility can play an important role. In this paper I analyze two specific questions relevant in this context. First, I explore if regional mobility is sensitive to regional unemployment rate differentials and look if regional mobility is higher in regions with higher regional unemployment rates. And second, I examine the destination choice of those who are regionally mobile. I investigate if they on average move towards regions with lower unemployment rates.

The analysis is based on a sample of job changes of men in West Germany between 1984 and 1997. I use register data, the IAB employment subsample.

The first question analyzed is if regional mobility is higher in regions with higher unemployment rates. My measure of regional mobility is the share of job changes which cover a geographical distance of more than 100 km. Indeed I find regional mobility to be sensitive to regional unemployment rates. It is somewhat higher if regional unemployment is higher. In a linear probability model for the share of regional changes among job changes I find regional mobility to be 0.1 percentage points higher when regional unemployment is 1 percentage point higher. When I include district specific effects the estimate is twice the size with 0.2 percentage points. However, compared to the average share of regional mobility of around 9 percent and also to personal covariates (especially recent regional mobility, education and employment status) the explanatory content of regional unemployment for regional mobility is rather small. Furthermore it is noteworthy that the sensitivity of regional mobility to regional unemployment is not homogeneous. Regarding age, only for the up to 40 year old regional mobility is higher when regional unemployment is higher. For older men I find no responsiveness. Comparing the two decades, the 80ies and the 90ies, the responsiveness is only present in the 80ies. No relationship between regional unemployment and regional mobility is found for the 90ies.

The second question I seek to answer is if those who are regionally mobile move towards regions with lower unemployment rates. This would be necessary in order to equalize unemployment rate differentials via regional mobility. Taking all regional changes (i.e. all job changes which cover a geographical distance of at least 100 km) I compare the unemployment rates in the origin region and in the destination region. I find that the unemployment rates in the destination regions are on average a third of a percentage point lower than in the origin regions in the eighties but less than 0.1 percentage point lower in the nineties. So those who are mobile on average move towards regions with lower unemployment rates. Comparing the 80ies and the 90ies this is much less the case in the nineties than in the eighties. A second notable difference between groups is, that job changers who experience unemployment between both jobs have a stronger tendency to move towards regions with lower unemployment than job changers who directly change jobs.

Of course, many open questions remain. How responsive is regional mobility after 1997? How responsive is it in East Germany? How do the results of the job change based analysis of regional mobility chosen in this paper relate to results based on individuals? What role plays regional unemployment in a structural model of internal migration compared to other economic factors and the household context?

3.5 References

Arntz, M. (2005). "The geographical mobility of unemployed workers. Evidence from West Germany". ZEW Discussion Paper No. 05-34.

Bender, S., A. Haas, and C. Klose (2000). "IAB employment subsample 1975–1995." *Schmollers Jahrbuch (Journal of Applied Social Science Studies)* 120, 649–662.

Büttner, T. (1999). "Agglomeration, Growth and Adjustment". Physica-Verlag, Heidelberg.

Decressin, J. W. and A. Fatás (1995). "Regional Labor Market Dynamics in Europe". *European Economic Review* 39(9), 1627-1655.

Dustmann, C. (2003). "Return Migration, Wage Differentials, and the Optimal Migration Duration". *European Economic Review* 47(2), 353-369.

Fitzenberger, B. and A. Spitz, (2004). "Die Anatomie des Berufswechsels: Eine empirische Bestandsaufnahme auf Basis der BIBB/IAB-Daten 1998/1999". ZEW Discussion Paper No. 04-05.

Fitzenberger, B., A. Osikominu, and R. Völter, (2006). "Imputation Rules to Improve the Education Variable in the IAB Employment Subsample". *Schmollers Jahrbuch (Journal of Applied Social Science Studies)* 126, 405–436.

Greenwood, M. (1997). "Internal Migration in Developed Countries" in Handbook of Population and Family Economics, Vol 1b. Edited by M. Rosenzweig and O. Stark. North Holland.

Gregg, P., S. Machin and A. Manning (2004). "Mobility and Joblessness" in Seeking a Premier League Economy, R. Blundell, D. Card and R. Freeman (eds.), University of Chicago Press.

Hatzius, J. (1994). "Regional Migration, Unemployment and Vacancies: Evidence from West German Microdata", Oxford Applied Economics Discussion Paper Series.

Herzog, H. W. Jr., A. M. Schlottmann and T. P. Boehm (1993). "Migration as Spatial Job-Search: A survey of Empirical Findings". *Regional Studies*, Vol. 27,4, 327-340.

Pissarides, C. and J. Wadsworth (1989). "Unemployment and the Inter-Regional Mobility of Labor". *The Economic Journal*, 99, 739–755.

Schündeln, M., (2007). “Are Immigrants More Mobile Than Natives? Evidence From Germany”. Discussion Paper, Harvard University.

Windzio, M., (2004). “Zwischen Nord- und Süddeutschland: Die Überwindung räumlicher Distanzen bei der Arbeitsmarktmobilität”. *Zeitschrift für Arbeitsmarktforschung*, 1/2004, 29-44.

Yankow, Jeffrey J., (2002). “The Geographic Mobility of Displaced Workers: Do Local Labor Market Conditions Matter?”. Discussion Paper, Furman University.

3.6 Appendix

Table 3.1: Job Change Sample Overview

	Frequency	Percent
Distance		
0 km	192,172	62.22
1–50 km	68,048	22.03
51–100 km	19,804	6.41
101–250 km	16,811	5.44
251 km	12,006	3.89
Interim period between old and new job		
Direct change without interruption	134,759	43.63
Unemplyoment	109,620	35.49
Nonemployment	64,462	20.87
Recalls	51,851	16.79
N	308,841	100.00
Persons	115,548	

Note: This table shows in its top part the geographical distance between the old job and the new job in km grouped into five distance categories. The sample can also be divided into direct changes, changes with unemployment and changes with nonemployment. Direct changes are job changes without interruption, in which the meantime between the old and the new job is less than 8 days. Job changes with an intervening period of at least 8 days are subdivided into changes with unemployment (at least one payment of unemployment benefits) and changes with nonemployment (no payment of unemployment benefits). Recalls are job changes in which the firm of the new job is the same as the one of the old job identified by a plant identifier.

Table 3.2: Summary statistics

Variable	Description of Variable	Mean	SD
move	Dummy for geographical distance between jobs of at least 100 km. Multiplied by 100.	9.331	29.086
distanzkm	Geographical distance between both jobs in km	34.020	84.897
alq	Regional unemployment rate at district level	9.028	3.477
alqwd	National unemployment rate	8.498	1.253
direct	Dummy. Direct job change	0.436	0.496
persunemp	Dummy. Job change with intervening unemployment	0.355	0.478
persnonemp	Dummy. Job change with intervening nonemployment	0.209	0.406
age2029	Dummy. $20 \leq Age \leq 29$	0.447	0.497
age3039	Dummy. $30 \leq Age \leq 39$	0.278	0.448
age4049	Dummy. $40 \leq Age \leq 49$	0.162	0.369
age5065	Dummy. $50 \leq Age \leq 65$	0.113	0.316
edu_l	No vocational training degree	0.161	0.368
edu_m	Vocational training degree	0.770	0.421
edu_h	University or Technical College degree	0.069	0.253
married	Married	0.410	0.492
foreign	Not German	0.123	0.329
stib0	Apprentice	0.055	0.227
stib1	Unskilled blue collar employee	0.310	0.462
stib4	White collar employee	0.225	0.417
earn	Real daily earnings in 1995 German Mark	124.551	52.313
earnsq	*earn* squared / 100	182.497	146.992
earncens	Dummy for earnings censored at social insurance threshold	0.044	0.204
tenure	Tenure in years	2.595	3.702
tenure1yr	Dummy. Tenure of more than one year	0.483	0.500
firmsize1_9	Dummy. Number of employees 1–9	0.238	0.426
firmsize10_99	Dummy. Number of employees 10–99	0.407	0.491
firmsize100_999	Dummy. Number of employees 100–999	0.239	0.426

Continued on next page...

... table 3.2 continued

firmsize1000plus	Dummy. Number of employees 1000+	0.117	0.321
mexp100y2	Dummy. At least one regional change in the last two years	0.071	0.256
rclexpy2	Dummy. At least one recall in the last two years	0.129	0.335
LEDexpy2	Dummy. At least one unemployment spell with payment of benefits in the last two years	0.294	0.456
urban	Dummy. City with more than 100,000 inhabitants.	0.369	0.483
year85		0.067	0.25
year86		0.072	0.258
year87		0.070	0.255
year88		0.071	0.258
year89		0.076	0.265
year90		0.081	0.272
year91		0.077	0.267
year92		0.072	0.259
year93		0.067	0.251
year94		0.070	0.256
year95		0.070	0.256
year96		0.071	0.257
year97		0.072	0.258

Note: The covariates can either relate to the old job or to the new one. Personal covariates (like age) relate the new job, i.e. are based on information reported by the new employer. Job related covariates (like earnings and firm size) relate to the old job. The unemployment rate in the district of the old job is measured at the time when the new job starts.

Table 3.3: Occupation Classification

No.	Designation	Groups	Mean	SD
1	Farming, Fishing, and Forestry Occupations	01–06	0.035	0.185
2	Miners, Extraction Workers	07–09	0.005	0.070
3	Building Material Production Workers, Ceramists	10–13	0.010	0.101
4	Chemical Workers and Plastic Workers	14–15	0.017	0.129
5	Paper- and Woodworkers, Printing Workers	16–18	0.016	0.126
6	Metal Workers	19–24	0.029	0.167
7	Mechanics	25–30	0.118	0.323
8	Electricians	31	0.044	0.205
9	Assembly Workers, Other Metal Workers	32	0.018	0.131
10	Textile and Apparel Workers	33–36	0.004	0.064
11	Leather Workers	37	0.002	0.044
12	Food Related Occupations	39–43	0.047	0.212
13	Construction Related Occupations	44–47	0.120	0.325
14	Plasterer, Insulation Workers, Tiler, Interior Decorators	48–49	0.019	0.135
15	Cabinetmakers	50	0.021	0.145
16	Painters	51	0.029	0.167
17	Machine Operators	54	0.012	0.111
18	Technicians	62–63	0.036	0.186
19	Retail and Wholesale Sales Workers	68	0.052	0.222
20	Financial Services Clerks, Forwarding Agents	69–70	0.025	0.157
21	Transportation Occupations	52, 71–74	0.139	0.346
22	Managers, Office and Administrative Support Occupations	75–78	0.076	0.264
23	Other Occupations	79–93	0.087	0.282
24	Engineers, Natural Scientists	60–61	0.021	0.144
25	No Occupation	98–99	0.018	0.132

Note: The two-digit occupations given in the IAB data (groups) were aggregated into 25 occupations. The classification is adopted from Fitzenberger and Spitz (2004). The table shows the means and standard deviations of the respective dummy variables.

Table 3.4: Industry Classification

No.	Designation	Groups	Mean	SD
1	Agriculture, Forestry, Fishing and Hunting	00–03	0.029	0.167
2	Utilities and Mining	04–08	0.012	0.108
3	Chemical Manufacturing	09–13	0.030	0.171
4	Nonmetallic Mineral Product Manufacturing	14–16	0.023	0.150
5	Primary Metal Manufacturing	17–22	0.028	0.166
6	Fabricated Metal Product Manufacturing, Machinery Manufacturing and Transportation Equipment Manufacturing	23–32	0.106	0.307
7	Electrical Equipment, Appliance, and Component Manufacturing	34–38	0.056	0.230
8	Wood Product and Furniture Manufacturing, Printing and Textiles	39–51	0.048	0.213
9	Apparel Manufacturing	52–53	0.003	0.056
10	Food, Beverage and Tobacco Product Manufacturing	54–58	0.031	0.173
11	Construction	59–61	0.183	0.387
12	Wholesale and Retail Trade	62	0.128	0.334
13	Transportation	63–68	0.081	0.273
14	Finance and Insurance	69	0.019	0.136
15	Services for Consumers	70–73	0.060	0.237
16	Education, Publishing, Health	74–78	0.039	0.193
17	Services for Businesses and Misc. Services	79–86	0.080	0.271
18	Organizations, Private Households and Public Administration	87–94	0.045	0.208

Note: The two-digit industries given in the IAB data (groups) were aggregated into 18 industries. The table shows the means and standard deviations of the respective dummy variables.

Table 3.5: Estimation Results: Regional Mobility without Covariates

	(1)	(2)	(3)	(4)	(5)
alq	0.018	–	0.046*	0.046*	0.154**
	(0.017)	–	(0.018)	(0.019)	(0.044)
alqwd	–	-0.141**	-0.192**		
	–	(0.047)	(0.051)	–	–
Year Dummies	–	–	–	yes	yes
District Dummies	–	–	–	–	yes
Intercept	9.167**	10.525**	10.545**	8.513**	14.027**
	(0.170)	(0.412)	(0.412)	(0.300)	(2.032)
N	308,841	308,841	308,841	308,841	308,841
R^2	0.000	0.000	0.000	0.000	0.013

Note: Linear probability model for the probability of regional mobility in percent. Dependent variable: *move* (dummy for regional mobility multiplied by 100). Robust standard errors clustered at the person level. †, ∗, and ∗∗ denote significance at the 10%, 5% and 1% level, respectively.

Table 3.6: Estimation Results: Regional Mobility with Covariates

District Dummies	not included		included	
Variable	Coeff.	(SE)	Coeff.	(SE)
alq	0.090**	(0.018)	0.199**	(0.041)
age3039	0.952**	(0.140)	0.943**	(0.140)
age4049	0.669**	(0.177)	0.647**	(0.177)
age5065	-0.594**	(0.191)	-0.638**	(0.191)
edu_l	-0.644**	(0.134)	-0.590**	(0.135)
edu_h	7.565**	(0.414)	7.457**	(0.412)
married	-1.033**	(0.123)	-1.058**	(0.123)
foreign	0.404*	(0.169)	0.424*	(0.173)
mexp100y2	33.279**	(0.361)	32.955**	(0.361)
rclexpy2	-1.066**	(0.151)	-1.344**	(0.157)
LEDexpy2	-2.954**	(0.133)	-2.928**	(0.133)
stib0	0.905**	(0.268)	0.888**	(0.267)
stib1	-0.455**	(0.124)	-0.428**	(0.125)
stib4	3.787**	(0.272)	3.749**	(0.271)
earn	-0.017**	(0.005)	-0.016**	(0.005)
earnsq	0.020**	(0.002)	0.021**	(0.002)
earncens	3.372**	(0.508)	3.361**	(0.507)
tenure	-0.274**	(0.015)	-0.274**	(0.015)
firmsize1_9	-2.897**	(0.157)	-2.806**	(0.158)
firmsize10_99	-1.948**	(0.142)	-1.917**	(0.142)
firmsize1000plus	-1.059**	(0.198)	-1.056**	(0.199)
urban	0.100	(0.125)		
year85	-0.211	(0.264)	-0.210	(0.263)
year86	0.140	(0.270)	0.205	(0.273)
year87	-0.055	(0.270)	0.012	(0.271)
year88	-0.660*	(0.266)	-0.580*	(0.267)
year89	-0.837**	(0.267)	-0.586*	(0.284)
year90	-0.877**	(0.271)	-0.467	(0.303)
year91	-1.010**	(0.276)	-0.541†	(0.326)
year92	-0.750**	(0.284)	-0.342	(0.324)

Continued on next page...

... table 3.6 continued

Variable	Coeff.	(SE)	Coeff.	(SE)
year93	-1.634**	(0.281)	-1.371**	(0.297)
year94	-1.741**	(0.272)	-1.579**	(0.281)
year95	-2.016**	(0.272)	-1.899**	(0.280)
year96	-2.242**	(0.271)	-2.225**	(0.273)
year97	-1.854**	(0.271)	-1.910**	(0.270)
Intercept	7.280**	(0.510)	6.075**	(0.646)
Occupation dummies	yes		yes	
Industry dummies	yes		yes	
District dummies	no		yes	
N	305,849		305,849	
R^2	0.1479		0.1517	

Note: Linear probability model for the probability of regional mobility in percent. Dependent variable: *move* (dummy for regional mobility multiplied by 100). Robust standard errors clustered at the person level. †, ∗, and ∗∗ denote significance at the 10%, 5% and 1% level, respectively.

Table 3.7: Estimation results: Regional Mobility with Interaction Effects

District Dummies	not included		included	
Variable	Coeff.	(SE)	Coeff.	(SE)
(1) Interaction with Age				
alq_age2029	0.124**	(0.023)	0.234**	(0.044)
alq_age3039	0.137**	(0.033)	0.259**	(0.050)
alq_age4049	-0.010	(0.038)	0.107*	(0.052)
alq_age5065	0.003	(0.038)	0.103*	(0.052)
age3039	0.817*	(0.356)	0.704*	(0.356)
age4049	1.885**	(0.428)	1.807**	(0.429)
age5065	0.500	(0.431)	0.553	(0.434)
...				
R^2	0.1480		0.1517	

Continued on next page...

... table 3.7 continued

Variable	Coeff.	(SE)	Coeff.	(SE)
(2) Interaction with Previous Mobility				
alq_mexp100y2	-0.136	(0.104)	-0.167	(0.105)
alq	0.100**	(0.018)	0.210**	(0.041)
mexp100y2	34.489**	(1.003)	34.435**	(1.003)
...				
R^2	0.1479		0.1517	
(3) Interaction with Tenure				
alq	0.037*	(0.019)	0.155**	(0.041)
alq_tenure1yr	0.112**	(0.014)	0.116**	(0.014)
tenure	-0.337**	(0.017)	-0.339**	(0.017)
...				
R^2	0.1481		0.1519	
(4) Interaction with Time				
alq_y8489	0.172**	(0.021)	0.186**	(0.041)
alq_y9097	-0.043	(0.027)	-0.001	(0.056)
...				
R^2	0.1480		0.1518	

Note: In the estimations the regional unemployment rate is interacted with the respective covariates. Otherwise they repeat the specification from table 3.6.

Table 3.8: Estimation results: Regional Mobility within Subgroups

District Dummies	not included		included	
Variable	Coeff.	(SE)	Coeff.	(SE)
(1) Only Direct Job Changes				
alq	0.099**	(0.031)	0.267**	(0.077)
...				
Intercept	6.581**	(1.021)	6.997*	(2.741)
N	133,797		133,797	
R^2	0.1597		0.1667	
(2) Only Job Changes with Unemployment				
alq	0.044†	(0.025)	0.161**	(0.053)
...				
Intercept	7.481**	(0.747)	12.357**	(2.979)
N	108,829		108,829	
R^2	0.1314		0.1360	
(3) Only Job Changes with Nonemployment				
alq	0.234**	(0.044)	0.404**	(0.117)
...				
Intercept	7.846**	(1.070)	11.928*	(5.295)
N	63,223		63,223	
R^2	0.1301		0.1374	

Note: The estimations repeat the estimations from table 3.6 for subgroups of job changes: direct job changes, job changes with an unemployment period and job changes with an nonemployment period.

Table 3.9: Summary Statistics: Regional Changes

Type of Change	All		Direct		Unemployment		Nonemployment	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
distanzkm	256.72	(136.06)	255.02	(132.15)	256.47	(140.59)	261.15	(140.22)
direct	0.54	(0.50)	1	(0)	0	(0)	0	(0)
persunemp	0.25	(0.43)	0	(0)	1	(0)	0	(0)
persnonemp	0.22	(0.41)	0	(0)	0	(0)	1	(0)
year8489	0.42	(0.49)	0.41	(0.49)	0.46	(0.50)	0.39	(0.49)
year9097	0.58	(0.49)	0.59	(0.49)	0.54	(0.50)	0.61	(0.49)
age3039	0.35	(0.48)	0.38	(0.49)	0.32	(0.47)	0.31	(0.46)
age4049	0.19	(0.39)	0.21	(0.41)	0.19	(0.39)	0.14	(0.34)
age5065	0.09	(0.29)	0.11	(0.32)	0.07	(0.26)	0.06	(0.23)
edu_l	0.09	(0.29)	0.06	(0.23)	0.11	(0.31)	0.16	(0.37)
edu_h	0.20	(0.40)	0.25	(0.43)	0.13	(0.33)	0.14	(0.35)
married	0.42	(0.49)	0.51	(0.50)	0.35	(0.48)	0.27	(0.45)
foreign	0.10	(0.30)	0.07	(0.25)	0.12	(0.32)	0.17	(0.38)
mexp100y2	0.32	(0.47)	0.31	(0.46)	0.32	(0.47)	0.35	(0.48)
rclexpy2	0.06	(0.23)	0.04	(0.19)	0.08	(0.27)	0.07	(0.26)
LEDexpy2	0.22	(0.41)	0.14	(0.35)	0.37	(0.48)	0.23	(0.42)
stib0	0.04	(0.20)	0.03	(0.16)	0.06	(0.23)	0.06	(0.24)
stib1	0.19	(0.39)	0.12	(0.32)	0.26	(0.44)	0.30	(0.46)
stib4	0.47	(0.50)	0.61	(0.49)	0.33	(0.47)	0.30	(0.46)
N	28,817		15,441		7,073		6,303	

Note: This table gives an overview about the regional changes, i.e. job changes with a geographical distance between the jobs of at least 100 km. It also provides summary statistics for the three subgroups of regional changes: direct job changes, job changes with an intervening period of unemployment and job changes with an intervening period of nonemployment.

Table 3.10: Estimation Results : Direction Index on Time Dummies

Type of Change	All	Direct	Unemployment	Nonemployment
Variable				
year8489	-0.362**	-0.297**	-0.614**	-0.188*
	(0.035)	(0.053)	(0.084)	(0.095)
year9097	-0.077**	-0.068*	-0.158**	-0.015
	(0.022)	(0.032)	(0.055)	(0.053)
N	28,817	15,441	7,073	6,303
R^2	0.001	0.001	0.002	0.000

Note: Regression of the direction index for regional changes on two dummies for the eighties and the nineties. Robust standard errors (clustered at the individual level) in parenthesis. Samples are in column 1 all regional changes and in columns 2–4 the three subgroups of regional changes.

Table 3.11: Estimation Results : Direction Index on Covariates

Type of Change	Direct		Unemployment		Nonemployment	
Variable	Coeff.	(SE)	Coeff.	(SE)	Coeff.	(SE)
year8489	-0.580**	(0.099)	-0.832**	(0.144)	-0.448**	(0.145)
year9097	-0.331**	(0.085)	-0.436**	(0.128)	-0.257*	(0.119)
age3039	-0.131	(0.080)	-0.070	(0.128)	0.115	(0.123)
age4049	0.069	(0.094)	0.184	(0.153)	0.139	(0.163)
age5065	-0.087	(0.107)	0.351†	(0.198)	-0.203	(0.217)
edu_l	-0.127	(0.138)	-0.223	(0.167)	-0.078	(0.146)
edu_h	0.112	(0.070)	0.012	(0.163)	-0.095	(0.165)
married	0.070	(0.069)	-0.037	(0.118)	0.075	(0.126)
foreign	-0.024	(0.125)	-0.248†	(0.148)	-0.113	(0.139)
mexp100y2	0.488**	(0.081)	0.877**	(0.131)	0.600**	(0.121)
rclexpy2	0.122	(0.178)	0.062	(0.218)	-0.120	(0.200)
LEDexpy2	-0.029	(0.102)	-0.429**	(0.130)	-0.026	(0.133)
stib0	0.145	(0.217)	-0.955**	(0.272)	-0.558*	(0.240)
stib1	0.205†	(0.122)	0.505**	(0.149)	0.254†	(0.140)
stib4	0.125	(0.076)	0.245†	(0.132)	-0.028	(0.143)
N	15,335		7,031		6,199	
R^2	0.004		0.017		0.008	

Note: Regression of the direction index for regional changes on two time dummies as in table 3.10 and additional covariates. Robust standard errors (clustered at the individual level) in parenthesis. Samples are the three subgroups of regional changes.

Chapter 4

Imputation Rules to Improve the Education Variable in the IAB Employment Subsample

4.1 Introduction

The IAB employment subsample (IABS) has become an important data source for empirical research on the German labor market. The IABS is a panel data set comprising administrative records for employment spells and for spells with transfer payments during periods of unemployment, see Bender *et al.* (2000). Compared to popular survey data sets like the German Socioeconomic Panel, the main advantages of the IABS are its large size, the long time period it covers, the almost complete absence of panel mortality, and the reliability of the core variables like date and length of spells, earnings, or type of transfer payments. However, it is well known that a number of variables are less reliable since they are not related to the purpose of the administrative reporting process producing the data. Nevertheless, research on the reliability of the IABS has been very scarce (see Fitzenberger, 1999, Steiner and Wagner, 1998, for rare exceptions). Earlier work (Cramer, 1985, or Schmähl and Fachinger, 1994) pointed to problems in administrative data on employment.

The returns to education and the skill bias in labor demand are two very important issues studied in labor economics (see e.g. Card, 1999, Katz and Autor, 1999, Fitzenberger, 1999) which require a reliable measure of formal education. The IABS contains the variable *BILD* comprising information on secondary and tertiary schooling degrees as well as on completion of a vocational training degree (apprenticeship). *BILD* is based on the reports by employers and the information is extrapolated to subsequent transfer spells. This education variable exhibits a number of apparent problems. First, there is missing information for 9.52% of the spells in the data set. Second, the education variable suffers from a large number of inconsistencies for a person over time. According to the reporting rule, employers are supposed to report the highest formal degree attained by the employee and not the degree required for the job. Hence, if a person is reported to have a certain educational degree and afterwards is reported to have a lower degree, we know that at least one of these reports must be wrong. We observe such inconsistent sequences of reports for 18.1% of the individuals in the data set. If the incidence of these problems is not completely at random, using the uncorrected data may result in misleading conclusions about the distribution of education and the relationship with other variables. Most of the empirical literature based on the IABS seems to use the uncorrected education variable and to exclude the observations with missings. Steiner and Wagner (1998) interpret missings in the education variable as saying that the employee exhibits no post secondary degree.

This paper develops various imputation procedures to improve the information in the IABS education variable. The main idea of our imputation approach is as follows: The panel nature of the data does not only allow us to identify inconsistencies but, under reasonable assumptions, it also allows us to deduce the likely education level of a person whose education information is missing or is inconsistent for a small number of spells. If the education information is missing for a small number of spells, we impute the likely education from past or future information. If a reported degree differs for a small number of spells from the likely education, we conclude that the currently reported education is incorrect and we impute the likely education instead. Imputation has been used before to improve the education variable in the IABS. This paper extends upon the earlier work of Fitzenberger (1999, appendix) and Bender *et al.* (2005, chapter 3.4). We develop a number of further refinements of the basic imputation procedures. Using different versions of our imputation procedure as benchmarks, we investigate the sensitivity of empirical estimates in typical applications.

Our deductive, nonstochastic imputation rule uses the available information in the data to develop a heuristic solution to the complex problems of missing data and misclassified reports. Based on hypotheses about the reporting behavior of employers, we impute logically correct values for the actual education level, which is basically time invariant (after an individual has reached his highest degree), when missings or inconsistencies occur. By using different hypotheses on the reporting behavior of the employers, we evaluate qualitatively the sensitivity of estimation results to the exact implementation of the imputation rule and we study the statistical nature of reporting errors. A statistical validation of our imputation methods lies beyond the scope of this paper.

There exist alternative approaches in the literature which use the misclassified data directly and take misclassification into account. These methods are application specific. Molinari (2004) makes exogenous assumptions about the misclassification probabilities and estimates identification regions for the true distribution based on the observed distribution of the misclassified data. Kane *et al.* (1999) estimate the returns to education when education is misclassified. The study relies on two measures of education which can both be mismeasured. These measures have to be (mean) independent of each other and of the wage conditional on the true education.

The latter assumption is not likely to hold in our context. For instance, as will be discussed in detail below, if the inconsistencies in the education variable are mainly due to underreporting the level of education for people who are overqualified with respect to their position, underreporting is associated with a low wage given true education. Lewbel (2003) gives necessary assumptions to estimate average treatment effects when treatment is misclassified. Again, these conditions are unlikely to hold in our context.

Multiple imputation methods are concerned with the problem of missing data.[1] Gartner and Rässler (2005) apply multiple imputation to the problem of censored wages in the IABS. The multiple imputation framework is appropriate in a situation where the underlying (non-)response process can be modeled using the observed data. For the following reasons, an application of these methods would be very difficult in the present context. First, the variable *BILD* does not only suffer from missing data but, in addition, there is also substantial misclassification in the non-missing data, for which the classical measurement error model is not appropriate. Second, even under the assumption that missing values are the only problem, it is questionable whether the basic condition, that the occurrence of missing values is random conditional on observed quantities, holds in the present context. In fact, if an employee is overqualified for his job, his employer may well underreport or not report at all the correct education of the employee because it does not match with the requirements and/or the social standing of the job. In the case of non-response, this would be correlated with the unobserved true education. Third, extensions of multiple imputation models to the panel data case can be computationally cumbersome and note that we have spell data with different numbers of spells per individual.

In light of the above discussion, our deductive imputation approach has three important advantages. First, we develop a heuristic and tractable solution to a complex problem where it is very difficult to apply existing methods. Second, our method is general in the sense that we do not rely on any (conditional) independence or distributional assumptions (see discussion above on multiple imputation method proposed in the literature). Third, using different versions of our imputation procedure as benchmarks, we can investigate the sensitivity of empirical estimates in typical applications. However, this does not allow us to assess directly the statistical variability induced by the imputation.

[1] An introduction and an overview over multiple imputation methods can be found in the textbooks of Little and Rubin (2002), Schafer (1997), or Rubin (1987).

A limitation of our approach is that it is not possible to tell which one of the different imputed education variables is the best. However, our results clearly suggest that it is advisable to use some correction of the education variable instead of ignoring the problem or resorting to *ad hoc* methods. As a practical rule, we recommend to conduct the analysis based on all the imputation procedures proposed in this paper and to check whether substantive results obtained are insensitive to the imputation procedure employed.[2]

The remainder of the paper is structured as follows. Section 4.2 describes the IABS data and provides details on the problems concerning the education variable. Section 4.3 develops the different imputation procedures to improve the education variable. Section 4.4 examines typical applications to compare the outcomes of the imputation procedures. Section 4.5 concludes. The appendix includes detailed results.

4.2 The IAB Employment Subsample (IABS)

4.2.1 Basic Description of IABS and *BILD*

We use the IABS version for the time period 1975-1997 distributed with detailed regional information, see Bender *et al.* (2000). Our imputation procedures are relevant for all versions of the IABS. The data contain daily register data for 589,825 individuals in Germany on employment spells and on spells with transfer payments from the Federal Labor Office (formerly *Bundesanstalt für Arbeit*). The IABS is a representative 1% sample of employment. After the end of the year and when a job ends, employers have to report earnings and other socio-demographic information about their employees like their educational degree. The earnings information and the length of the employment spells are used to calculate contributions to and benefits from the social insurance system and, hence, are very reliable. Periods of self-employment and employment as life-time civil servants (Beamte) that are not subject to (mandatory) social insurance are not included in the data.

The education information has to be reported with every employment spell but it bears no relevance for the social security system. To our knowledge, reporting the

[2]For the case of nonresponse and censoring, identification of bounds on population parameters also avoids untestable assumptions about the distribution of the missing data, see Horowitz and Manski (1998, 2000). This method is useful for analyzing 'worst-case' scenarios. It could be fruitful to explore this in future work as an alternative to our imputation approach.

employee's education incorrectly has no consequences. This explains why the education variable *BILD* in the IABS is less reliable compared to information on earnings or the beginning and ending of spells. Spells on transfer payments and technical spells documenting gaps in the employment history, for instance, due to military service or maternity leave, do not provide new information on the educational level. Instead *BILD* is extrapolated during such spells based on the information in the most recent employment spell. Thus, we base our imputation procedures only on the information given in employment spells. On average, the data contain 14.6 spells per person of which 12.3 are employment spells.

Since the variable *BILD* is based on employer reports to the social security system, it is an important question how the reporting system changed between 1975 and 1997, possibly affecting the reported education. As mentioned before, the basis of the IAB employment subsample is the integrated reporting system for the social insurance, i. e. the statutory health, pension and long-term care insurance. The notification procedure was introduced in the former Federal Republic of Germany on 1 January 1973 and on 1 January 1991 – after the German reunification – in the new Länder and Berlin-East, too. Since 1973 there have been several revisions of the legislation governing the formal way of how the notifications have to be submitted by the employers.[3] However, these changes did not concern the content – for instance the precision – of the demographic variables contained in the so called "Tätigkeitsschlüssel".[4] Thus, we conclude that inconsistencies in the education variable over time are in fact attributable to employers' unreliability and not to institutional changes. This is supported by the finding that the probabilities of inconsistent and missing education reports show only very slight changes over the years (tables 4.2 and 4.5).

The education information in the IABS distinguishes four different educational de-

[3]A first major revision took place in 1981 when the "Zweite Datenerfassungsverordnung" and the "Zweite Datenübermittlungsverordnung" came into effect. Their main goal was to improve the completeness of the overall amount of notifications in order to provide correct aggregate employment statistics (Wermter and Cramer, 1988). Second, in 1984 there was a change in the scope of gross earnings which are subject to social security contributions ("Änderungsverordnung zur 2. DEVO" (Bender *et al.*, 1996, and Fitzenberger, 1999, appendix)). A third major revision of the notification procedure came into effect in 1999 ("Datenerfassungs- und übermittlungsverordnung"). Now, all employers are required to provide uniform information which is automatically processed.

[4]The "Tätigkeitsschlüssel" comprises variables that describe the job content and the qualification of the employee (cf. http://www.arbeitsagentur.de/content/de_DE/hauptstelle/a-07/importierter_inhalt/pdf/schluessel.pdf).
The new "Datenerfassungs- und übermittlungsverordnung" actually intended to introduce a new "Tätigkeitsschlüssel" which though has not yet been implemented so far.

grees ($\equiv$ successful completion): high school (*Abitur*), vocational training, technical college (*Fachhochschule*), and university. University is considered the highest degree, a technical college the second highest. Since there is no clear ranking between high school and vocational training, employers have to choose among all four combinations between the two. Thus, *BILD* can take six possible meaningful values:

1. no degree at all (henceforth: ND),
2. vocational training degree (VT),
3. high school degree (HS),
4. high school degree and vocational training degree (HSVT),[5]
5. technical college degree (TC), and
6. university degree (UD).

We argue that these six educational outcomes can be ranked in increasing order except that no ranking exists between the second degree VT and the third degree HS. We consider the comprehensive degree HSVT to be higher than both HS and VT. Furthermore, if the employee's education is not known, it can be reported as missing. According to the reporting rule, employers are supposed to report the highest degree attained by the employee, not the degree required for the current job. As a consequence, the sequence of education records should be non-decreasing over time because one can only attain higher degrees over time, not lose them. A decreasing sequence violates the reporting rule and represents evidence for inconsistent reporting behavior. All imputation procedures developed in this paper provide a corrected education variable with consistent information over time.

4.2.2 Spells with Missing Education

Table 4.1 (in the appendix) reports the distribution of the variable *BILD* in the original data. As can be seen, 9.52% of the spells exhibit missing education information. One might suspect that missing values are mostly a problem concerning non-employment spells and short employment spells. Therefore, we also calculate

[5] In the following, we will refer to HSVT as if it is a separate degree even though it is in fact a combination of two degrees.

the distribution of the education variable among full time working males in 1995 excluding apprentices and weighting the spells by their length. The weighting changes the unit of measurement from spells to person years to correspond to employment. Still a weighted share of 7.35% has missing education information. Therefore, missing values are also a sizeable problem among employees.

Next, we investigate how the incidence of missing education information among employees is related to other observed covariates in the IABS. We estimate a probit modeling the probability of a missing education report as a function of personal characteristics.[6] The estimation is based on employment spells only. Table 4.2 displays the marginal effects on the probability of a missing report. Most of the effects are significant but they are not very large compared to an observed rate of 7.8% of missing information. Noteworthy are a 6.1 (SE 0.1) percentage points (ppoints) higher probability of a missing report for foreigners relative to Germans, a 9.2 ppoints (SE 0.3) higher probability for part-time workers with less than half the regular hours compared to fulltime salaried employees and considerable differences in reporting quality across industries. Compared to the investment goods industry, the probability of a missing report is 15.7 (SE 0.4) ppoints higher in consumer services and 10.8 (SE 0.3) ppoints higher in the main construction trade.

4.2.3 Changes in Education across Spells

Compared to missing information, changes in the education information reported across spells are more difficult to deal with and it is crucial to analyze the sequence of reported education records across spells. If first a high degree and afterwards a low degree is reported, we know that this sequence is inconsistent with the reporting rule, but we do not know which report is incorrect. It can be the first one overreporting or the second one underreporting or even both can be incorrect. However, we can identify whether an entire sequence is consistent, i.e. nondecreasing. In the sample, 81.9% of the persons exhibit consistent sequences of education information while 18.10% do not.

[6]We thank Alexandra Spitz for providing a useful classification for occupations based on the *Alphabetisches Verzeichnis der Berufsbenennungen der Bundesanstalt für Arbeit* (cf. the manual accompanying the IABS). We adjusted the classification to the occupation information given in the regional file (*BERUF*=1-117) as follows: (i) farmers/farm managers *BERUF*=1, 2 (ii) service workers *BERUF*=97, 98, 110-116 (iii) operatives/craft *BERUF*=3-57, 78-85 (iv) sales workers *BERUF*=70-73 (v) clerical workers *BERUF*=74-77, 89-96 (vi) administrative, professional and technical workers *BERUF*=58-69, 86-88, 99-109.

Example 1: Person with inconsistently reported education

SPELL	*BILD*	Education	Employer	Employed
1	1	ND	1	yes
2	1	ND	2	yes
3	5	TC	3	yes
4	1	ND	1	yes
5	1	ND	0	unemployed
6	1	ND	4	yes
7	2	VT	5	yes
8	2	VT	6	yes
9	2	VT	0	unemployed
10	2	VT	0	unemployed
11	-9	missing	7	yes
12	2	VT	8	yes
13	2	VT	0	unemployed
14	2	VT	9	yes
15	1	ND	10	yes
16	2	VT	9	yes

Example 1 shows a hypothetical but representative person (all examples in this paper show hypothetical cases) with inconsistently reported education records. Only spell 3 shows education TC but all later spells show lower education with ND or VT or missing education. We do not know if the report of TC is true, but the decrease in reported education afterwards shows that some report violates the rule of reporting the highest attained degree. Either the report of TC itself is wrong, or, in fact, the employee obtained the degree after spell 2 and before spell 3. In the latter unlikely case, all education reports after spell 3 showing a lower degree would be incorrect. In this example, there exists a second inconsistency. The report of ND at spell 15 is lower than the report of VT at spell 14.

Some insights on the reporting behavior of employers can be gained by looking at consecutive pairs of education records for the same employee. Overall, in 91.5% of all cases, two consecutive reports are the same.[7] But there is a sharp difference depending on which employer issued the report. If both reports are by the same employer, they coincide in 97.0% of the cases. However, if issued by two different employers, this rate amounts to only 63.2%. The higher stability of reports by the same employer is to be expected for the following reasons. First, attaining a higher degree often coincides with changing the employer. The second explanation

[7]The descriptive statistics in this section are based on employment spells only.

is rather technical and is related to the artificial splitting of some employment spells in the IABS in order to assure data privacy. This results in two consecutive spells by the same employer with the same education information. Third, this may also indicate that employers just replicate their previous reports causing serial correlation of reporting errors for reports on the same employee.

Furthermore, we investigate the conditional probabilities for reported education conditional on the previous report for a given person. Such a transition matrix is calculated for reports by the same employer in table 4.3 and for reports from differing employers in table 4.4. The high numbers (above 93% except for HS) on the diagonal in table 4.3 confirm that the same employer is likely to repeat the report given before. When the reporting employer changes, VT still has a probability of 76.9% to be repeated in the next record. The probability for UD to be repeated is 74.7%. This is not surprising since VT and UD are likely to be the highest degrees people attain. The other educational outcomes are, on the contrary, reported in a less stable way with probabilities of being repeated reaching at most 55.1%.

In table 4.5, we estimate the probability that consecutive pairs of education reports on the same employee are inconsistent, i. e. the second report is lower than the first one. Since with an inconsistent pair we do not know whether the first or the second report is wrong, but only that at least one of them must be incorrect, we consider reported characteristics both in the first and the second spell. The covariates describing the employment status have the largest coefficients. Working as a trainee (apprentice obtaining a VT) at the second spell of the consecutive pair increases the probability of an inconsistent pair by 4.9 ppoints (SE 0.09), relative to working as a salaried employee. This compares to an observed total rate of 2.1% for all pairs. Working as a skilled worker at the first spell of the pair leads to a 3.4 ppoints (SE 0.06) higher probability of an inconsistent pair, again compared to working as as salaried employee. Industry and nationality only weakly affect the probability of inconsistent reports. This is in sharp contrast to the influence that these variables have on the probability of a missing report (see table 4.2).

4.3 Imputation Procedures

This section develops three imputation procedures. All imputation procedures are based on extrapolation of degrees which we will describe first.

4.3.1 Extrapolation and Reporting Errors

Extrapolation of educational degrees is based on three facts:

(i) the formal education level of an individual can increase, when an additional degree is attained, or stay constant but it can not decline,

(ii) the formal education of an individual remains mostly constant once the individual has entered working life, and

(iii) employers have to report the highest attained degree.

Facts (i) and (ii) state that the education of individuals is monotonically increasing but mostly constant. Fact (iii) ensures that the employers have to report the actual education of their employees and not the education necessary for the particular job which might be lower. Thus the reported education has to be equal to the actual education and hence also to be monotonically increasing.

Extrapolating plausible education reports to later spells with lower or missing education reports, we can construct an improved education variable which is monotonically increasing. For the extrapolation of education, it is helpful to distinguish three types of reporting errors: (i) underreported education, (ii) not reported education resulting in missings, and (iii) overreported education. Spells with underreported or not reported education can be imputed with the correct education if one extrapolates the correct education from an earlier spell. In contrast, overreported education can not be corrected by extrapolation of a correctly reported degree because the overreported degree is higher. Even worse, if one extrapolates an overreported degree to later spells with correctly reported education the quality of the education variable deteriorates. This is because extrapolation has a ratchet effect. After extrapolation, the imputed education will monotonically increase but not go down.

The obvious challenge for every imputation rule is to detect spells which are likely overreports and not to extrapolate the information. Since we do not have exogenous information about the true education but only know the reported education we cannot compare the results of an imputation rule with the true information. Hence evaluation criteria for imputation procedures requiring the true values to be known, like those in Chambers (2001), are also not applicable.[8] Instead, we propose three

[8]Other evaluation criteria exist in the literature, see e.g. Rubin (1996).

imputation procedures bounding the true education in distribution from above and below. The first imputation procedure (IP1) extrapolates the highest education level ever observed, including overreports. Thus, IP1 can be viewed as an upper bound for the true education. The other two procedures IP2 and IP3 are more conservative by only extrapolating reliable reports, thus resulting in lower bounds for the true education level. IP2 uses the frequency of the report of a specific degree as an indication of its reliability and only extrapolates degrees which are reported at least three times. IP3 assesses the reporting quality of employers and only extrapolates reports from reliable employers. Put together, these imputation procedures provide benchmarks reflecting the range of the true education information. If substantive results do not differ between IP1 and IP2 or IP3, we argue that they basically coincide with results obtained for a correct measure of education. Then, it also seems justifiable that standard errors for the estimated quantities are not adjusted for the remaining uncertainty inherent in the imputation. The next subsections will give details of the imputation procedures.[9]

4.3.2 Imputation Procedure 1 (IP1)

The first imputation procedure IP1 puts no restrictions in extrapolating degrees. Every education report and hence also every overreport can be extrapolated. We argue that this procedure is likely to impute the correct education or to overstate the true education. Since we observe several education reports per person it is quite likely the true education will be eventually reported or an overreport occurs. In these cases the true education will be imputed or even an overreport will be imputed to many spells due to the ratchet effect. It is also possible that this imputation procedure understates the true education for some persons. An example is the case where the true education level is never reported.

The imputation procedures are implemented in four steps. Step 1 defines which reports can be extrapolated – either all as in IP1 or only reports deemed reliable as in IP2 or IP3. The procedures only differ in this first step. The following three steps contain the actual extrapolation and further adjustments. Next we describe the details.

Step 1: Preparation for Extrapolation

[9] More details can be found in Fitzenberger, Osikominu, and Völter (2005).

Step 1 distinguishes valid spells with extrapolatable information and invalid other spells. IP1 uses all employer reports for extrapolation. The nonemployment spells in the IABS (benefit payment spells, interruption spells) do not carry independent education information but repeat the education information of the most recent employment spell. Hence we do not use this information but treat the spells as spells with missing information. Steps 2 and 3 will extrapolate information to these invalid spells as to other spells with missing information. The original data include educational degrees for persons below the age of 18 years which often seem implausible. Therefore, we first impute ND for all spells in this age range.

Step 2: Forward Extrapolation

Step 2 implements the extrapolation of degrees to later spells. The procedure goes over all spells of an individual person starting with the first spell and ending with the last one. It extrapolates degrees to later spells if the education information reported in these later spells is lower or missing. The extrapolation stops if a spell with a higher degree or the last spell of the person is reached.

The extrapolation of education information to subsequent spells has to account for the fact that the degrees HS and VT cannot be ranked. When persons have both degrees this has to be explicitly reported. Hence, the extrapolation rule imputes HSVT if it reaches a spell with one of the two degrees and there is another previous spell for this person with the other degree.

Step 3: Backward Extrapolation

The forward extrapolation in step 2 leaves the education information missing, when spells with missing values precede a person's first spell with a valid educational report. Since the educational degree of a person is rather time constant, we also extrapolate backwards the first valid educational degree to previous spells with missing information. We do not extrapolate backwards degrees beyond degree specific age limits, because the attainment of a certain degree implies a certain amount of years of schooling. The age limits are the median ages at which the degrees are reported for the first time for persons in the data. We do not impute backwards UD below the age of 29 years, TC below 27 years, HSVT below 23 years, HS below 21 years, and VT below 20 years. If the first information reported is ND, this is imputed to all spells before. Note that the first spells of young persons can comprise missing education values, even if these persons show non-missings values in subsequent spells.

Step 4: Additional adjustments

For persons with missing education information in all spells, we impute VT if their employment status is skilled worker (*Facharbeiter*), foreman (*Polier*) or master craftsman (*Meister*). This is justified by the fact that in almost 90% of the cases with valid education information we observe the degree VT together with such an employment status. Therefore, we impute VT for those cases. Subsequently, we also extrapolate the imputed information VT forwards and backwards analogous to steps 2 and 3.

If persons only have employment spells with other information on employment status and education is missing for all spells, we leave it at that.

The data contain a number of parallel spells for persons who hold two or more jobs at the same time. If the imputed education variable so far takes different values for parallel spells we finally impute the highest education information among the parallel spells to these parallel spells.

Example 2 illustrates the implementation of IP1. The forward extrapolation (*Step 2*) extrapolates VT from spell 3 to spells 4 and 5, where the lower education level ND is reported. At spell 7, HS is reported. With VT having been reported before, we assume both degrees, HS and VT, are held and impute HSVT. HSVT is considered higher than HS and extrapolated to spell 8 and 9. For spell 10, UD is reported. Even though it is reported only once for this person, IP1 extrapolates UD to spells 11 and 12 because IP1 extrapolates every degree. Forward extrapolation alone would leave spells 1 to 2 with missing information. Hence (*Step 3*), we extrapolate backwards VT from spell 3 to spells 1 to 2. It can be seen that the imputed sequence is consistent (i.e. non-decreasing), which by construction is the case for all imputed data. In example 2, there is no missing information left. This is not necessarily the case, especially when there is only missing information about a person.

4.3.3 Imputation Procedure 2 (IP2)

Imputation procedure 2 (IP2) is a conservative imputation procedure, which is likely to understate the true education by restricting extrapolation to degrees which are reported at least three times. The frequency of a report serves as a measure of

Example 2: IP1

SPELL	*BILD*	Education	IP1	IP1 Education
1	-9	missing	2	VT
2	-9	missing	2	VT
3	2	VT	2	VT
4	1	ND	2	VT
5	1	ND	2	VT
6	2	VT	2	VT
7	3	HS	4	HSVT
8	3	HS	4	HSVT
9	3	HS	4	HSVT
10	6	UD	6	UD
11	2	VT	6	UD
12	2	VT	6	UD

its reliability. If a degree is reported repeatedly, then we assume it has a lower probability to be an overreport than if is reported only once or twice. There are occasions in which a low frequency of a report arises quite naturally without indicating a likely overreport, e.g. when two degrees are obtained within a short time period. Therefore, we implement procedure 2 in two versions, IP2A and IP2B. Procedure IP2A strictly restricts extrapolation to degrees which are at least three times reported. Procedure IP2B is less strict. Only when an inconsistent sequence of education reports indicates reporting errors for a person, extrapolation is restricted to degrees which are reported at least three times. If a person's education sequence is consistent, then IP2B extrapolates every report just as IP1.

Compared to IP1, IP2A and IP2B yield possibly lower imputed educational reports since not all reported higher degrees are extrapolated. By construction, imputed values for IP2B lie between IP1 and IP2A.

IP2A and IP2B proceed by the same four steps as described above for IP1. The only differences involve step 1:

In step 1, IP2A accepts all employment spells as valid for extrapolation when the reported degree is reported in at least three spells.[10] When counting specific education reports, we only count employment spells, not non-employment spells. This reflects the fact that only information at employment spells is directly employer reported.

[10]If the total number of employment spells for a person is only four, the minimum frequency for acceptance is reduced to two reports. If there are less than four employment spells, educational information in every employment spell is treated as valid.

Example 3: IP2A

SPELL	*BILD*	Education	Employed	Frequency of report	IP2A	IP2A Education
1	1	ND	yes	7	1	ND
2	1	ND	yes		1	ND
3	5	TC	yes	1	1	ND
4	1	ND	yes		1	ND
5	1	ND	yes		1	ND
6	2	VT	yes	3	2	VT
7	2	VT	unemployed		2	VT
8	2	VT	yes		2	VT
9	2	VT	yes		2	VT
10	-9	missing	yes		2	VT
11	1	ND	yes		2	VT
12	1	ND	yes		2	VT
13	1	ND	yes		2	VT

Only information from valid spells will be extrapolated later. Spells with invalid information are later treated as spells with missing information, meaning information will be extrapolated to them from valid spells. Analogous to IP1, nonemployment spells are treated as invalid spells.

IP2B uses the same heuristic rule for extrapolation as IP2A but only for persons with an inconsistent sequence of educational reports in the original data. For persons with a consistent sequence, all employment spells are accepted as a basis for extrapolation.

Both IP2A and IP2B do not accept degrees for persons below the age of 18 but impute ND instead. For young persons below the age of 23 years in vocational training the educational information ND or HS is accepted even without being reported frequently enough.

Example 3 illustrates the implementation of IP2A.

First, it is determined which degrees are reported at least three times to be valid for extrapolation. ND is reported seven times and hence valid. TC is only reported once. Thus, spell 3 will be treated as a spell with missing information. VT is reported in three employment spells and hence also considered as valid. Actually, it is reported four times in the example, but spell 7 is an unemployment spell just repeating information from spell 6. Hence, spell 7 is treated as a spell with missing information. Now, extrapolation can proceed. ND is extrapolated from spell 2 to spell 3 previously containing the invalid TC report. VT is extrapolated from spell 9 to spell 10 with missing information and spells 11-13 with the lower report of ND. Since the reported information in this example is inconsistent, IP2B would proceed the same way.

4.3.4 Imputation Procedure 3 (IP3)

Analogous to IP2, IP3 is designed as a conservative imputation procedure, likely understating the true education. We do not take the frequency of a report as a sign of its reliability but try to judge the reporting quality of the reporting employer.

We consider the reporting quality of an employer as being good, when he always reports the same education for an employee or changes his report only once.[11] We only extrapolate reports from good reporters. IP3 treats reports from bad reporters as missing and extrapolates reports from good reporters to these spells. The hypothesis underlying this procedure is the following. Employers do not reevaluate the educational degree of their employees every time they have to give a report but tend to copy from previous reports. Thus, the frequency of the reports as such is not that informative. It is more informative if an employer changes his report about an employee. And since it is very unlikely that persons attain two (or more) new degrees while being employed with one employer we think two (or more) changes in the report indicate bad reporting quality. As noted above, education is constant for most workers after entering the labor market.

Employers might change their reports in order to correct previous reporting errors. IP3 tries to explicitly take this into account. We allow for two types of self correction. The first type consists of errors corrected immediately: an employer changes the reported degree only for one spell and switches back immediately afterwards. In this case, we ignore the switch back–and–forth and the employer is still classified as reliable. The second type of self correction concerns reliable employers. If they inconsistently change their report from a higher degree to a lower degree, we assume they always wanted to report the lower degree. If a reliable employer permanently changes to a higher degree, we interpret this as the actual attainment of the higher degree.

[11]Note that our data do not allow us to identify whether different employees are employed by the same employer. We can only identify which of a person's employment spells are with the same employer. Hence the reporting quality is in fact match specific.

Now, we start extrapolating a degree when it is reported for the first time by a reliable employer unless the employer did not intend to report this higher degree as a consequence of self correction. In the latter case, we first impute the intended lower degree which becomes the basis for extrapolation.

Again, compared to IP1, IP3 yields potentially lower imputed educational reports since not all reported higher degrees are extrapolated. In this respect, one cannot rank IP3 relative to IP2A or IP2B.

Furthermore, IP3 proceeds by the same four steps as described above for IP1, IP2A, and IP2B. The only differences regard step 1.

For a given individual, IP3 preliminarily accepts all employment spells with non-missing education information which are the first reports of a given employer and, in addition, all employment spells by the same employer whenever the reported degree changes. When a change occurs, the type of reporting error is classified. If there is an immediate self correction, we impute the intended report in the deviating spell. Then, if we count not more than one change in the reported degree, the respective employer is classified as reliable. Otherwise, the employer is classified as unreliable and his reports are set to missing. Next, inconsistent reports by reliable employers are corrected in the spells which were first accepted. As in all other procedures, we impute spells from persons below the age of 18 with ND.

From this point onwards, extrapolation proceeds as for the other imputation procedures.

Example 4 illustrates the implementation of IP3.

Before extrapolation can take place the reliability of the employers has to be determined and self corrected reporting errors have to be detected. Employer 1 changes the reported information once and is hence reliable. Employer 2 changes the reported information twice, thus he is not reliable. His spells (4-6) are treated as spells with missing information. Employer 3 seems to change the reported level of education three times. But we interpret the report of HSVT at spell 8 as an immediate correction of a one time misreport because VT is reported by this employer at spells 7 and 9. Hence, we count only one change and classify this employer as reliable. We conclude that he intended to report VT at spells 7-10 and HSVT at spell 11. Employer 4 changes the reported degree once and is hence reliable. But the report of TC after UD is inconsistent. IP3 assumes this to be a self correction

Example 4: IP3

SPELL	*BILD*	Education	Employer	Employer reliable	Intended report	IP3	IP3 Education
1	1	ND	1	yes	1	1	ND
2	1	ND	1		1	1	ND
3	2	VT	1		2	2	VT
4	2	VT	2	no		2	VT
5	3	HS	2			2	VT
6	4	HSVT	2			2	VT
7	2	VT	3	yes	2	2	VT
8	4	HSVT	3		2	2	VT
9	2	VT	3		2	2	VT
10	2	VT	3		2	2	VT
11	4	HSVT	3		4	4	HSVT
12	6	UD	4	yes	5	5	TC
13	5	TC	4		5	5	TC
14	5	TC	4		5	5	TC
15	5	TC	4		5	5	TC
16	4	HSVT	5	yes	4	5	TC
17	4	HSVT	5		4	5	TC

and that employer 4 always intended to report TC. Employer 5 never changes the reported education and is classified as reliable. Now, extrapolation can take place on the basis of the reliable employer's intended reports (i. e. after taking account of the self correction of reporting errors at spell 8 and 12). VT is extrapolated from spell 3 to spells 4-6. TC is extrapolated from spell 15 to spells 16 and 17.

4.4 Empirical Analysis

This section compares the corrected education data resulting from the different imputation procedures to the original data. We study (i) the education mix in employment, (ii) wage inequality between and within education skill groups, (iii) earnings regressions and how misreports are related to earnings, and (iv) the incidence of underreports.

Our basic imputation approach is based on plausible assumptions about the reporting behavior of employers and the previous section shows the importance of missing values and inconsistencies in the education variable. Therefore, we believe that empirical results using the imputed education variable are more reliable than using the uncorrected data. If substantive empirical results do not differ considerably (measured by the economic importance of the changes) when applying the different imputation procedures, we argue that the results coincide with results obtained for a correct measure of education. If results do differ, then we cannot provide a point estimate for the quantity of interest and we suspect that the different estimates provide bounds for the point estimate based on the correct education data. We cannot account further for the statistical uncertainty inherent in our imputation.

4.4.1 Education Mix in Employment

Table 4.6 shows the education shares for the original data and for the respective imputed data resulting from procedures IP1, IP2A, IP2B, and IP3 where the shares have been calculated based on the raw spells, i. e. all unweighted spells. To assess the relevance of the imputation procedure for practical applications, table 4.7 reports education shares for men working fulltime in 1995 in West Germany weighted by the spell length. The tables show that all procedures could eliminate most of the missing values. Their share decreases from 9.5% to 1.9-3.2% of the raw spells. Considering the weighted sample, we see a similar picture at a lower level. The share of missing values decreases from 7.4% to 1.2-2.1%. The remaining missing values can be explained by two reasons: (i) persons with all education information missing and (ii) the age limits for backwards extrapolation of degrees. The imputation procedures do not only reduce the share of missing information but also the share of ND and HS. The shares of the education groups VT, HSVT, TC, and UD increase for the raw spells as well as for the weighted data. Next we discuss the results for the weighted data in more detail. The by far largest increase in absolute terms concerns the category VT with an increase of 5.6-6.9 ppoints (added to 65.3% initially). HSVT shows the largest increase in relative terms: +1.1-2.4 ppoints (added to 2.7% initially). Considering the higher education levels, UD gains more (+0.8-1.3 ppoints added to 5.0% initially) than TC: 0.4-0.7 ppoints added to 4.0% initially. The decrease in ND is 2.5-5.0 ppoints from 15.1%. The size and change of HS is small.

The imputation procedures decrease the shares of ND and HS and result in a higher educational attainment among employed workers. The share of the employees holding any degree is higher (lower share of ND) and the share of the higher educational levels (TC, UD) is higher.

Comparing the different imputation procedures, IP1 shows the strongest impact on the educational composition. IP1 results in the highest shares of the higher education categories (HSVT, TC, UD) which is to be expected since it potentially extrapolates any higher report. IP2A changes the educational composition least strongly. The resulting shares of the high education categories (HSVT, TC, UD) are the lowest. IP2B is comparable to procedure IP2A except for a lower share of missing information and a higher share of VT. IP3 gives shares which are roughly in the middle between procedure IP1 and procedure IP2A. This shows that our acceptance rules based on frequency are stricter than the acceptance rule based on the reliability of the employer.

Are the differences between the imputed data from the procedures small compared to the difference to the original data? This would imply that it is important to use an imputation rule, but not that important which one. Certainly the differences between the different imputed data are small concerning the missing values and VT. For the other categories the differences are also not too large except for the small category of HSVT. Its share goes up from 2.7% to between 3.8% (IP2A) and 5.1% (IP1). For this category, the differences between the procedures are not negligible.

Additional insights on how the different procedures work can be gained from looking at the conditional imputation probabilities for the different education categories given the reported education. These imputation matrices are reported in tables 4.8 to 4.11 based on the raw spells because the procedures are based on unweighted employment spells. The tables are transformation matrices in which the diagonal elements give the probability an original report remains unchanged by the imputation procedure and the off diagonal elements in each row give the probability it is imputed with one of the other education categories. All procedures impute spells containing missing information with ND in about 25% of all cases and with VT in about 50% of all cases. The large values on the diagonals in the tables show that all procedures leave at least 73% of the non-missing reports unchanged. Reports from the largest category VT are rarely changed, with the procedures leaving more than 95.6% unchanged. Only UD reports are changed less often, more than 97.1% of them are unchanged. HS reports exhibit the highest rate of being imputed with other information. They remain unchanged with a rate of only 73.0-77.1% and, if changed, they are most likely to be imputed with HSVT in 9.9-18.9% of the cases. ND reports are quite likely to be changed, too. 77.4-83.7% of them are unchanged. 15.7-21.3% are imputed with VT. Even if the broad picture looks similar and all procedures provide an upward correction of the education variable, there are differences between the procedures. Only IP1's entries below the diagonal are all close to zero. This reflects the fact that every reported non-missing degree is used for extrapolation under the assumption of no overreporting. Procedures IP2 and IP3, on the contrary, do not extrapolate every degree but some reported degrees classified as unreliable are imputed with lower degrees. This mainly concerns HS which is in 1.3-4.4% of the spells imputed with ND and in 4.5-6.1% with VT. This also concerns HSVT reports. Those reports stand a 3.3-5.7% chance of being imputed with VT.

In the literature, the six educational categories are often aggregated into three groups: (U) without a vocational training degree [ND and HS], (M) with a vocational training degree [VT and HSVT], and (H) with a higher educational degree [TC and UD] (see for instance Fitzenberger, 1999). This makes imputations within these groups irrelevant but a considerable number of imputations takes place across the groups U, M, and H, like the imputation of VT to ND. Hence, the imputation procedures are relevant at the more aggregated level as well. But the aggregation reduces the differences concerning the educational distribution, since the small group HSVT with the largest differences is aggregated with VT.

4.4.2 Wage Inequality Between and Within Education Groups

Now, we investigate the impact of the imputation procedures on measures of wage inequality between and within skill groups. For illustrative purposes, we focus on wage inequality among men working full time in West Germany and only consider two years, 1984 and 1997. We aggregate the six education categories into three skill groups, U, M, and H, as described in the last subsection. Table 4.12 shows the 20th, the 50th and the 80th percentile of the daily wage (in German Marks/DEM) for men in 1984 and in 1997 by the skill groups U, M and H. For the high skilled H, the 50th and the 80th percentiles cannot be calculated since wages are right censored in the data at the social security threshold. The table shows that the percentiles of the daily wage estimated with the imputed data are in most cases some DEM lower compared to those calculated with the original data. In 1984, this only concerns the wage percentiles for the skill groups M and H, which are estimated 1 to 4 DEM lower with the imputed data than with the original data (originally 90-143 DEM). In 1997, this concerns all skill groups, the estimated daily wage percentiles are up to 9 DEM lower for the imputed data. Lower estimated wage percentiles resulting from the imputed data are consistent with our view that there are many more underreports than overreports and underreports are associated with employees holding degrees which employers do not consider necessary for the job. Therefore, employees with underreports earn less than employees holding the same, correctly reported degree.

As a measure of wage inequality between skill groups, we consider the difference in log daily wages between the skill groups M and U at the 50th and the 20th percentiles and between the skill groups H and M only at the 20th wage percentile due to censoring of the higher wages. The numbers are given in table 4.13. Imputing the education variable has a noticeable influence on the estimates of wage inequality between the low skilled U and the medium skilled M at the 20th wage percentile. In 1984, the estimate is lower with 0.095 for all procedures instead of 0.118 with the original variable, whereas in 1997 the estimate is higher, i. e. 0.168 to 0.219 compared to 0.163. Since the differences go in the opposite direction the estimated 1984-97 increase in wage inequality varies considerably from 0.045 to 0.073-0.124. With other words, one would underestimate the average yearly growth rate of between wage inequality by a factor of at least two, i. e. amounting to 0.3% based on the original variable compared to 0.6-0.9% using the corrected education variables. Note further the large differences between the imputed variables themselves in the 1997 estimates which is translated to the trend estimates. The estimated inequality measures between U and M at the 50th and between H and M at the 20th wage percentile and the respective trends are not changed in a systematic way, however, there are noticeable differences as well.

Table 4.14 reports wage inequality within the skill groups U and M. It shows the differences in log wages between the 80th and the 50th wage percentiles as well as the differences between the 50th and the 20th percentiles. Overall, the largest impact of the imputation procedures on measured inequality can be found in 1997 for skill group U below the median: the 50%-20% log wage difference is measured as 0.257-0.269 instead of 0.228 where the results of the different imputation procedures are quite close. The other measured within group wage inequalities are changed less than half that much, at most by 0.014. Concerning the trend between 1984 and 1997, the largest change can also be observed for the 50%-20% log wage difference for skill group U. Whereas the original data result in an increase of 0.036, the imputed show a larger increase of 0.065-0.086. The measured trend for U and M above the median is almost not affected by the imputation procedures: the growth in the 80%-50%

log wage difference for M shows a slightly smaller value with 0.003-0.014 compared to 0.021 for the original data.

Summing up, imputation affects some measures of wage inequality, especially in the lower part of the wage distribution of the low to medium skilled groups.

4.4.3 Mincer-type Earnings Regressions

This subsection investigates the effects of imputing education on estimated wage regressions. Furthermore, we investigate whether and how the measurement error in education is related to wages. We estimate the following Mincer-type earnings equation (Mincer, 1974):

$$\begin{aligned}\log w &= \alpha + \beta_{ND} d_{ND} + \beta_{HS} d_{HS} + \beta_{HSVT} d_{HSVT} \\ &\quad + \beta_{TC} d_{TC} + \beta_{UN} d_{UN} + \beta X + \varepsilon\end{aligned}$$

where $\log w$ is the log daily wage for fulltime employed men in West Germany in 1995. The education variables are dummies for five categories, the most frequent category VT is the omitted category. We also control for age and age squared in X. Since wages are censored from above at the social security threshold, we estimate a Tobit model. The results without further controls are given in table 4.15. It can be seen that the coefficients for ND, TC, and UN do not differ much between the original data and the different versions of the imputed data. The differences are below 8.7% of the coefficient obtained based on the original data. But the differences are partly significant due to small standard errors.[12] The intercept as a measure for the VT log wage also does not seem to differ significantly. But the coefficients for the smaller education categories HS and HSVT change. The difference is largest for HSVT between the original data and the data based on IP1, the coefficients are 0.196 (SE 0.007)[13] and 0.123 (0.005), respectively. IP3 gives results comparable to IP1 and IP2 gives results between IP1 and the original data. The coefficient for HS does not differ significantly due to large standard errors.

[12]We do not estimate the sampling variance of the difference when applying the different imputation procedures. However, if the sample variance of the coefficient is small in all cases then the variance of the difference is small because of the Cauchy-Schwarz-Inequality.

[13]Here and in the following standard error in parentheses.

We have repeated the estimations with additional controls for being foreign, six occupations and 13 industries (see table 4.16). The picture remains qualitatively the same. The additional controls reduce the coefficients in absolute value for the log wage differences. Again the coefficients for ND, TC, and UN are quite comparable. The intercept does not differ notably. The coefficients on HS differ but not significantly. The coefficients on HSVT again differ significantly. It is 0.086 (0.006) for the original data and differs most strongly for IP1 with 0.035 (0.005).

Regarding these regressions, the impact of correcting the education variable is fairly small. This is somewhat surprising since in tables 4.8-4.11 the composition of the individual education categories differs between the original data and the imputed data. For instance, with probability 21.3%, IP1 classifies a spell with reported education ND as VT. It is not surprising that the imputation procedures do not change the intercepts as estimates for the log wage for VT given the large size of this group and the low rate of change for spells reporting VT initially. The difference is largest for HSVT which is the education category affected most by the imputation procedures.

Next, we analyze the relationship between the individual wage and the incidence as well as the type of misreport estimating a wage regression. This is of importance since wage estimations in the spirit of Kane *et al.* (1999), which take misclassification explicitly into account, require conditional (mean) independence of wages and measurement error given true education. We can explore whether this assumption is likely to hold by assuming true education to be close to one of the corrected values. Then, we construct a missing dummy, which is one if the education information in the original data is missing, an overreport dummy, which is one if the report in the original data is higher, and an analogous underreport dummy. If measurement error is independent of the wage conditional on true education, the dummies for the measurement error types should have insignificant coefficients in the wage regressions with the improved data. The regression controls for being foreign, occupation and industry since the incidence of missing education information was shown above to be correlated with some of these variables.

The results can be found in table 4.17. The coefficient for a missing report varies between -0.106 (0.005) and -0.120 (0.004). The coefficients for underreported education are of similar size with values between -0.104 (0.004) and -0.116 (0.003). The coefficient on overreported education is significantly negative for IP2A with -0.181

(0.012) and IP2B with -0.211 (0.017) but insignificant for IP3. Since procedure IP1 assumes there are no overreports for persons over 17 years there is no coefficient to estimate. If we are willing to assume that the true education is not too far from one of the imputed education variables, we can conclude that an underreported or not reported education is associated with a 10% lower wage given the true education. A lower wage, when education is underreported, is in accordance with the hypothesis that some employers report the education required for the job, not the degree attained by the employee. They pay a wage corresponding to the lower reported education. The evidence on overreported education is not conclusive. Altogether it seems that measurement error and wage are not conditionally independent given the true education. Therefore, potential alternatives to imputation suggested in the literature are not applicable here.

4.4.4 Underreports and Overreports

This section returns to the question of incorrect education reports. Comparing the imputed data and the original data for employment spells in West Germany, the share of underreports lies between 5.8% (IP2A) and 8.8% (IP1) and the share of overreports between 0.2% (IP1) and 1.0% (IP2A). Underreports are quantitatively as important as missing values, whereas overreports are much less frequent. For this reason and because overreports differ by construction according to the different imputation procedures, we focus on underreports in the following.

The incidence of underreports is analyzed by comparing the reported education to the imputed education from IP2A in a probit regression with the set of regressors also used when analyzing missing education reports (see table 4.2). The marginal effects are reported in table 4.18. As the largest effect, we find a 5.7 ppoints (0.1) higher probability of an underreport for a non-skilled worker compared to a salaried employee. If the report comes from an employer who gives only one or two reports about this employee, the probability of an underreport is 3.7 ppoints (0.1) higher than when the employer gives more than five reports. Possibly, employers who anticipate employing a person only for a short time spend less effort reporting correctly. The effect of working in the main construction trade is also quite large, with a 2.6 ppoints (0.2) higher probability than in the investment goods industry. Note that, for the probability of a missing report, the effect of this industry is four times as large (see table 4.2) and, analyzing inconsistencies in table 4.5, there are almost no industry effects. In contrast to what we found for missing reports, foreigners are less likely to have underreports. The results for the other imputation procedures (not displayed here) are quite comparable.

4.5 Conclusion

The education variable in the IABS shows two apparent shortcomings: missing data and observed data which is inconsistent with the reporting rule for the variable. Based on the notion that the education variable should represent a person's highest degree, that the educational degree of a working person is rather time constant, and that people can only attain degrees over time, but not lose them, we propose different procedures to improve the variable by deductive imputation. There is no exogenous information to validate our imputation procedures. Using plausible hypotheses about the reporting process, we argue that our basic imputation procedure is likely to overstate true education and our two other refinements are likely to understate true education. If empirical results based on the different procedures are close, we argue that the imputed education variable is basically correct. In order to evaluate the impact of imputing the education variable, we analyze the educational distribution of employment as well as wage inequality between and within skill groups, and we compare wage regressions using the original data and the corrected data.

Imputation removes more than two thirds of the missing values. The corrected data are by construction consistent with the reporting rule. Concerning the education distribution of employment, the improvement of the data matters. All procedures give higher shares for vocational training (with or without a high school degree), technical college, and university degrees and lower shares for no degree or high school only. The resulting shares do not differ a lot in size between the procedures, except for the small category vocational training plus high school. In most dimensions, misreporting educational degrees especially affects wage inequality measured at lower percentiles in the low and medium skilled groups. We find, for instance, that for the unskilled the measured growth from 1984 to 1997 in the difference between the median wage and the 20th wage percentile is considerably higher compared to the original data. In Mincer-type wage regressions, improving the data makes only a small difference, except again for the small category vocational training plus high school. However, wages for three skill groups are typically lower at all degrees when using the imputed data.

Overall, our results indicate that there is some evidence in favor of the hypothesis that underreporting of educational degrees is a more severe problem than overreporting. In fact, employers tend to report the degree required for the position rather than the highest formal qualification attained by the employee. Moreover, our findings imply that usual *ad hoc* methods of dealing with or even ignoring the data quality issues regarding the education variable may bias results, especially if the focus of the analysis lies on subpopulations where the incidence of these problems is not negligible. We have demonstrated that, by exploiting the available information in the data as well as on the institutional context, it is possible to put so much structure on the problem as to recover, in a heuristic way, an education variable that is likely to be very close to the truth.

Our analysis does not provide a definite rule on how to choose among the different imputation procedures. However, we recommend using some correction of the education variable, as suggested in our paper, instead of the common practice in existing studies involving the use of an inconsistent education variable with a large number of missing observations. In addition, without correction, the education variable in the IABS tends to understate the educational level of the employees. In actual applications, where the education variable is crucial, we recommend to use all imputation procedures suggested here. If the substantive results obtained are insensitive to the use of imputation procedure, then it is very likely that they are not affected by the remaining uncertainty about the education variable. Clearly, more research on improving the data quality in the IABS is strongly needed, in particular, since the same database is used in recent evaluations of labor market reforms in Germany.

4.6 References

Bender, S., J. Hilzendegen, G. Rohwer, H. Rudolph (1996). "Die IAB-Beschäftigtenstichprobe 1975-1990", *Beiträge zur Arbeitsmarkt- und Berufsforschung*, 197, IAB, Nürnberg.

Bender, S., A. Haas, and C. Klose (2000). "IAB Employment Subsample 1975-1995", *Schmollers Jahrbuch*, 120, 649-662.

Bender, S., A. Bergemann, B. Fitzenberger, M. Lechner, R. Miquel, S. Speckesser, C. Wunsch (2005). "Über die Wirksamkeit von Fortbildungs- und Umschulungsmaßnahmen – Ein Evaluationsversuch mit prozessproduzierten Daten aus dem IAB", *Beiträge zur Arbeitsmarkt- und Berufsforschung*, 289, IAB, Nürnberg.

Card, D. (1999). "The Causal Effect of Education on Earnings", in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics Volume 3*, Amsterdam: Elsevier.

Chambers, R. (2001). "Evaluation Criteria for Statistical Editing and Imputation", *National Statistics Methodological Series No. 28.*

Cramer, U. (1985). "Probleme der Genauigkeit der Beschäftigtenstatistik", *Allgemeines Statistisches Archiv*, 69, 56–68.

Fitzenberger, B. (1999). *Wages and Employment Across Skill Groups: An Analysis for West Germany*, Heidelberg: Physica Verlag.

Fitzenberger, B., A. Osikominu, and R. Völter (2005). "Imputation Rules to Improve the Education Variable in the IAB Employment Subsample", ZEW Discussion Paper 05–10, Mannheim.

Gartner, H. and S. Rässler (2005). "Analyzing the Changing Gender Wage Gap Based on Multiply Imputed Right Censored Wages", IAB Discussion Paper 05/2000, Nürnberg.

Horowitz, J. L. and C. F. Manski (1998). "Censoring of Outcomes and Regressors Due to Survey Nonresponse: Identification and Estimation Using Weights and Imputations", *Journal of Econometrics*, 84, 37–58.

Horowitz, J. L. and C. F. Manski (2000). "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data", *Journal of the American Statistical Association*, 95, 77–84.

Kane, T. J., C. E. Rouse, and D. Staiger (1999). "Estimating Returns to Schooling when Schooling is Misreported", NBER Working Paper 7235.

Katz, L. and D. Autor (1999). "Changes in the Wage Structure and Earnings Inequality", in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics Volume 3*, Amsterdam: Elsevier.

Lewbel, A. (2003). "Estimation of Average Treatment Effects with Misclassification", Working Paper, Boston College.

Little, R.J.A. and D.B. Rubin (2002). *Statistical Analysis with Missing Data*, 2nd edition, New York: Wiley & Sons.

Mincer, J. (1974). "Schooling, Experience, and Earnings", National Bureau of Economic Research, New York.

Molinari, F. (2004). "Partial Identification of Probability Distributions with Misclassified Data", Working Paper, Cornell University.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*, New York: Wiley & Sons.

Rubin, D. B. (1996). "Multiple Imputation After 18+ Years", *Journal of the American Statistical Association*, 91, 473-489.

Schafer, J. L. (1997): *Analysis of Incomplete Multivariate Data*, London: Chapman & Hall.

Schmähl, W. and U. Fachinger (1994). "Prozeßproduzierte Daten als Grundlage für sozial- und verteilungspolitische Analysen – Erfahrungen mit Daten der Rentenversicherungsträger für Längschnittsanalysen", in: Hauser, R., N. Ott, G. Wagner (eds.), *Mikroanalytische Grundlagen der Gesellschaftspolitik*, Band 2, *Erhebungsverfahren, Analysemethoden und Mikrosimulation*, Berlin: Akademie Verlag.

Steiner, W. and K. Wagner (1998). "Has Earnings Inequality in Germany Changed in the 1980's?", *Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 118(1), 29-59.

Wermter, W. and U. Cramer (1988): "Wie hoch war der Beschäftigtenanstieg seit 1983?", *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 4, Institut für Arbeitsmarkt- und Berufsforschung, 468-482.

4.7 Appendix

Table 4.1: **Distribution of the Education Variable *BILD* in the Original Data**

Education (abbreviation)[a]	Coded as	Number of spells	Share of spells	Weighted share[b] male empl 1995
Missing	-9	819,701	9.52	7.35
No vocational training degree, no high school degree (ND)	1	2,325,379	27.00	15.13
Only vocational training degree, no high school degree (VT)	2	4,794,512	55.66	65.28
Only high school degree, no vocational training degree (HS)	3	95,955	1.11	0.59
High school degree and vocational training degree (HSVT)	4	153,728	1.78	2.69
Technical college degree (TC)	5	175,603	2.04	3.97
University degree (UD)	6	249,180	2.89	4.98
Total		8,614,058	100.00	100.00

Notes: [a] In German vocational training degree means *abgeschlossene Berufsausbildung*, high school degree *Abitur*, technical college degree *Fachhochschulabschluss* and university degree *Hochschulabschluss.* [b] Weighted Share describes the education reported for fulltime working males in West Germany in 1995. Apprentices are not included. Employment spells are weighted by their length.

Table 4.2: **Probit Regression of Education Information Missing**

Regressors	Marg eff	Robust SE	Regressors	Marg eff	Robust SE
≤19 years	-0.018	(0.001)**	spell≤30 days	0.023	(0.001)**
30-39 years	0.014	(0.001)**	30<spell≤180 days	0.015	(0.000)**
40-49 years	0.019	(0.001)**	1-2 reports by empl	0.038	(0.001)**
50-59 years	0.021	(0.001)**	3-5 reports by empl	0.023	(0.001)**
60+ years	0.024	(0.002)**	year 75	0.009	(0.001)**
female	-0.002	(0.001)**	year 76	0.008	(0.001)**
married	-0.009	(0.001)**	year 77	0.005	(0.001)**
foreign	0.061	(0.001)**	year 78	0.005	(0.001)**
trainee	-0.039	(0.001)**	year 79	0.005	(0.001)**
non-skilled worker	0.040	(0.001)**	year 80	0.002	(0.001)**
skilled worker	-0.023	(0.001)**	year 81	0.001	(0.001)
master craftsman/foreman	-0.034	(0.002)**	year 82	0.001	(0.001)
home worker	0.129	(0.012)**	year 83	0.000	(0.001)
part time ≤18h	0.092	(0.003)**	year 84	0.000	(0.000)
part time >18h	0.028	(0.001)**	year 86	-0.000	(0.000)
farmers/farm managers	0.011	(0.002)**	year 87	0.000	(0.001)
service workers	0.019	(0.001)**	year 88	0.001	(0.001)
sales workers	-0.016	(0.001)**	year 89	0.002	(0.001)**
clerical workers	-0.028	(0.001)**	year 90	0.005	(0.001)**
admin/profes/techn staff	-0.020	(0.001)**	year 91	0.007	(0.001)**
agriculture	0.014	(0.003)**	year 92	0.008	(0.001)**
basic industry	0.015	(0.002)**	year 93	0.010	(0.001)**
consumer goods industry	0.018	(0.002)**	year 94	0.012	(0.001)**
food industry	0.045	(0.002)**	year 95	0.013	(0.001)**
main construction trade	0.108	(0.003)**	year 96	0.013	(0.001)**
construction completion trade	0.054	(0.003)**	year 97	0.015	(0.001)**
trade	0.060	(0.002)**			
transport and communication	0.084	(0.003)**			
business services	0.090	(0.002)**			
consumer services	0.157	(0.004)**			
education, non profit org	0.022	(0.002)**			
public administration	0.026	(0.002)**			
Observed prob	0.078				
Predicted prob at $\bar{x}$	0.057				
N	6,369,039				
Pseudo R^2	0.123				

Notes: Dependent variable: dummy for reported education missing. Estimation based on all employment spells in West Germany. Base category: 20-29 years, male, not married, German, working fulltime as a salaried employee, occupation group operatives/craft, investment goods industry, more than five reports by the employer about the employee, spell longer than 180 days, 1985. Intercept included in estimation. Robust standard errors with clustering at the person level. * significant at 5%, ** significant at 1%.

Table 4.3: **Conditional Probabilities of Education Reported Given Previous Report by the Same Employer**

Education reported previously	Education reported later by the same employer						
	Missing	ND	VT	HS	HSVT	TC	UD
Missing	94.43	1.68	3.47	0.06	0.12	0.11	0.13
ND	0.35	93.73	5.76	0.06	0.07	0.02	0.01
VT	0.26	0.69	98.86	0.04	0.07	0.05	0.03
HS	0.36	1.31	4.38	87.25	5.74	0.42	0.55
HSVT	0.34	0.35	2.26	0.21	96.15	0.32	0.37
TC	0.17	0.10	0.98	0.06	0.20	98.18	0.31
UD	0.15	0.05	0.46	0.06	0.11	0.23	98.94
Total	6.87	25.51	59.27	0.98	1.90	2.36	3.10

Notes: The table contains the conditional probabilities that the row education will be reported for a person given the previous report for the person was the column education and was reported by the same employer. Based on all employment spells.

Table 4.4: **Conditional Probabilities of Education Reported Given Previous Report by a Different Employer**

Education reported previously	Education reported later by different employer						
	Missing	ND	VT	HS	HSVT	TC	UD
Missing	35.65	25.27	34.96	0.91	1.14	0.84	1.23
ND	12.48	53.12	31.62	1.16	0.70	0.46	0.47
VT	8.75	10.81	76.91	0.49	1.45	0.92	0.68
HS	7.92	15.94	23.12	25.61	10.52	5.10	11.80
HSVT	7.45	4.19	35.13	2.23	37.91	5.57	7.51
TC	5.70	1.89	21.77	1.05	5.35	55.08	9.17
UD	4.66	0.82	9.46	1.01	4.11	5.28	74.66
Total	12.84	24.34	54.71	1.19	2.08	1.88	2.96

Notes: The table contains the conditional probabilities that the row education will be reported for a person given the previous report for the person was the column education and was reported by a different employer. Based on all employment spells.

Table 4.5: **Probit Regression of Inconsistent Reports**

Regressors	Marg. Effect	Robust SE	Regressors	Marg. Effect	Robust SE
≤19 years	-0.0071	(0.0001)**	food industry	-0.0006	(0.0003)
30-39 years	0.0005	(0.0001)**	main construction trade	0.0062	(0.0004)**
40-49 years	-0.0007	(0.0001)**	construction completion trade	0.0012	(0.0004)**
50-59 years	-0.0025	(0.0001)**	trade	-0.0002	(0.0002)
60+ years	-0.0049	(0.0002)**	transport and communication	-0.0043	(0.0002)**
female	-0.0002	(0.0001)	business services	0.0019	(0.0003)**
married	-0.0026	(0.0001)**	consumer services	-0.0001	(0.0003)
foreign	-0.0013	(0.0001)**	education, non profit org	0.0031	(0.0003)**
spell≤30 days	0.0059	(0.0002)**	public administration	0.0025	(0.0004)**
30<spell≤180 days	0.0103	(0.0001)**	agriculture (t-1)	-0.0021	(0.0004)**
spell≤30 days (t-1)	0.0087	(0.0002)**	basic industry (t-1)	-0.0010	(0.0003)**
30<spell≤180 days (t-1)	0.0085	(0.0001)**	consumer goods industry (t-1)	0.0004	(0.0003)
1-2 reports by empl	0.0065	(0.0002)**	food industry (t-1)	0.0008	(0.0003)*
3-5 reports by empl	-0.0004	(0.0001)*	main construction trade (t-1)	-0.0013	(0.0003)**
1-2 reports by empl (t-1)	0.0214	(0.0003)**	construction compl trade (t-1)	-0.0009	(0.0003)**
3-5 reports by empl (t-1)	0.0094	(0.0002)**	trade (t-1)	0.0018	(0.0002)**
trainee	0.0486	(0.0009)**	transport and comm (t-1)	0.0092	(0.0005)**
non-skilled worker	0.0331	(0.0006)**	business services (t-1)	0.0012	(0.0003)**
skilled worker	-0.0154	(0.0001)**	consumer services (t-1)	0.0023	(0.0004)**
master craftsman/foreman	-0.0089	(0.0001)**	education, non profit org (t-1)	-0.0036	(0.0002)**
home worker	0.0278	(0.0034)**	public administration (t-1)	-0.0038	(0.0002)**
part time ≤18h	0.0227	(0.0008)**	year 75	0.0014	(0.0005)**
part time >18h	0.0096	(0.0004)**	year 76	-0.0007	(0.0003)**
trainee (t-1)	-0.0098	(0.0001)**	year 77	0.0015	(0.0003)**
non-skilled worker (t-1)	-0.0109	(0.0001)**	year 78	0.0030	(0.0003)**
skilled worker (t-1)	0.0338	(0.0006)**	year 79	0.0019	(0.0003)**
master craftsman/f (t-1)	0.0225	(0.0015)**	year 80	0.0026	(0.0003)**
home worker (t-1)	-0.0049	(0.0007)**	year 81	0.0022	(0.0003)**
part time ≤18h (t-1)	-0.0032	(0.0002)**	year 82	0.0020	(0.0003)**
part time >18h (t-1)	-0.0041	(0.0002)**	year 83	0.0003	(0.0003)
farmers/farm managers	0.0021	(0.0006)**	year 84	0.0001	(0.0003)
service workers	0.0023	(0.0003)**	year 86	0.0002	(0.0003)
sales workers	-0.0056	(0.0002)**	year 87	-0.0003	(0.0003)
clerical workers	-0.0037	(0.0002)**	year 88	-0.0004	(0.0003)
admin/profes/techn staff	-0.0071	(0.0002)**	year 89	-0.0005	(0.0002)*
farmers/farm man (t-1)	0.0031	(0.0006)**	year 90	0.0003	(0.0003)
service workers (t-1)	-0.0009	(0.0002)**	year 91	0.0002	(0.0003)
sales workers (t-1)	0.0112	(0.0005)**	year 92	-0.0007	(0.0002)**
clerical workers (t-1)	0.0078	(0.0004)**	year 93	-0.0006	(0.0002)*
admin/profes/techn (t-1)	0.0191	(0.0005)**	year 94	-0.0011	(0.0002)**
agriculture	0.0006	(0.0005)	year 95	-0.0012	(0.0002)**
basic industry	0.0010	(0.0003)**	year 96	-0.0025	(0.0002)**
consumer goods industry	0.0017	(0.0003)**	year 97	-0.0038	(0.0002)**
Observed prob	0.0213		N		5,474,652
Predicted prob at $\bar{x}$	0.0106		Pseudo R^2		0.2092

Notes: Dependent variable: dummy for reported education lower than in the previous report. (t-1) indicates variables concerning the previous employment spell. Estimation based on all employment spells in West Germany. Base category: 20-29 years, male, not married, German, working fulltime as a salaried employee, occupation group operatives/craft, investment goods industry, more than five reports by the employer about the employee, spell longer than 180 days, 1985. Intercept included in estimation. Robust standard errors with clustering at the person level. * significant at 5%, ** significant at 1%.

Table 4.6: **Distribution of Education Variable after Imputation, Unweighted Spells**

Education	Original data	IP1	IP2A	IP2B	IP3
Missing	9.52	1.90	3.10	2.09	3.24
No vocational training degree, no high school degree	27.00	23.41	25.68	25.80	24.09
Only vocational training degree, no high school degree	55.66	63.78	62.13	62.89	62.77
Only high school degree, no vocational training degree	1.11	1.07	1.03	1.06	1.03
High school degree and vocational training degree	1.78	3.63	2.47	2.54	2.99
Technical college degree	2.04	2.61	2.30	2.32	2.45
University degree	2.89	3.60	3.28	3.30	3.43
Total	100.00	100.00	100.00	100.00	100.00

Notes: Shares based on all 8,614,058 spells.

Table 4.7: **Distribution of Education Variable after Imputation, Weighted Male Employment in 1995**

Education	Original data	IP1	IP2A	IP2B	IP3
Missing	7.35	1.18	2.09	1.28	2.01
No vocational training degree, no high school degree	15.13	10.14	12.61	12.52	11.14
Only vocational training degree, no high school degree	65.28	72.15	70.83	71.59	71.60
Only high school degree, no vocational training degree	0.59	0.46	0.54	0.55	0.52
High school degree and vocational training degree	2.69	5.05	3.76	3.83	4.21
Technical college degree	3.97	4.71	4.37	4.40	4.49
University degree	4.98	6.32	5.81	5.84	6.03
Total	100.00	100.00	100.00	100.00	100.00

Notes: The table describes the education mix for men in West Germany working fulltime in 1995. Apprentices are not included. Spells are weighted by their length.

Table 4.8: **Imputation Matrix for Procedure IP1**

	Imputed data							
Original data	Missing	ND	VT	HS	HSVT	TC	UD	Total
Missing	19.61	24.77	49.30	0.87	2.50	1.38	1.57	100.00
ND	0.04	77.41	21.29	0.47	0.58	0.12	0.09	100.00
VT	0.03	0.27	95.83	0.00	2.56	0.83	0.47	100.00
HS	0.03	0.64	0.04	76.84	18.91	1.26	2.29	100.00
HSVT	0.02	0.10	0.04	0.01	89.64	5.04	5.14	100.00
TC	0.03	0.04	0.03	0.00	0.00	92.39	7.50	100.00
UD	0.04	0.02	0.01	0.01	0.01	0.00	99.92	100.00
Total	1.90	23.41	63.78	1.07	3.63	2.61	3.60	100.00

Notes: The table contains the conditional probabilities the column information will be imputed given the spell originally contains the row information. Based on all 8,614,058 spells.

Table 4.9: **Imputation Matrix for Procedure IP2A**

	Imputed data							
Original data	Missing	ND	VT	HS	HSVT	TC	UD	Total
Missing	28.39	24.98	42.45	0.68	1.25	1.01	1.22	100.00
ND	0.53	82.79	16.17	0.26	0.16	0.05	0.04	100.00
VT	0.30	1.53	96.10	0.06	1.25	0.46	0.29	100.00
HS	4.56	4.37	4.53	73.82	10.11	0.91	1.70	100.00
HSVT	0.92	1.88	5.67	1.79	82.88	3.29	3.56	100.00
TC	0.56	0.69	2.61	0.24	0.54	90.48	4.88	100.00
UD	0.42	0.29	1.09	0.25	0.36	0.49	97.10	100.00
Total	3.10	25.68	62.13	1.03	2.47	2.30	3.28	100.00

Notes: The table contains the conditional probabilities the column information will be imputed given the spell originally contains the row information. Based on all 8,614,058 spells.

Table 4.10: **Imputation Matrix for Procedure IP2B**

	Imputed data							
Original data	Missing	ND	VT	HS	HSVT	TC	UD	Total
Missing	20.43	27.79	46.99	0.84	1.51	1.12	1.33	100.00
ND	0.17	83.66	15.68	0.25	0.16	0.05	0.04	100.00
VT	0.10	0.89	97.00	0.06	1.23	0.45	0.27	100.00
HS	2.10	3.87	4.53	77.12	9.94	0.86	1.58	100.00
HSVT	0.42	1.30	4.46	0.88	86.31	3.18	3.43	100.00
TC	0.29	0.46	2.07	0.16	0.44	91.84	4.74	100.00
UD	0.22	0.19	0.83	0.15	0.30	0.37	97.93	100.00
Total	2.09	25.80	62.89	1.06	2.54	2.32	3.30	100.00

Notes: The table contains the conditional probabilities the column information will be imputed given the spell originally contains the row information. Based on all 8,614,058 spells.

Table 4.11: **Imputation Matrix for Procedure IP3**

	Imputed data							
Original data	Missing	ND	VT	HS	HSVT	TC	UD	Total
Missing	20.94	24.22	48.93	0.87	2.23	1.30	1.51	100.00
ND	2.53	78.97	17.55	0.39	0.41	0.09	0.06	100.00
VT	0.83	0.78	95.57	0.04	1.76	0.65	0.37	100.00
HS	2.37	1.32	6.09	72.97	14.31	1.08	1.86	100.00
HSVT	2.26	0.56	3.31	0.33	84.87	4.13	4.53	100.00
TC	1.14	0.26	1.74	0.13	0.51	90.21	6.01	100.00
UD	0.51	0.06	0.52	0.09	0.24	0.46	98.11	100.00
Total	3.24	24.09	62.77	1.03	2.99	2.45	3.43	100.00

Notes: The table contains the conditional probabilities the column information will be imputed given the spell originally contains the row information. Based on all 8,614,058 spells.

Table 4.12: **Wage Percentiles for Men by Skill Group for 1984 and 1997**

Year	Skill group	Percentile	Orig. data	IP1	IP2A	IP2B	IP3
1984	U	20	80	80	80	80	80
		50	97	96	97	97	97
		80	116	115	116	116	117
	M	20	90	88	88	88	88
		50	110	109	109	109	109
		80	145	142	143	143	142
	H	20	143	139	141	141	140
1997	U	20	113	107	106	108	109
		50	142	140	138	140	141
		80	173	170	169	171	173
	M	20	133	129	132	130	129
		50	166	161	164	163	162
		80	222	213	217	215	214
	H	20	206	197	207	203	198

Notes: The table contains the percentiles of the daily wages in DEM for men working fulltime in West Germany without apprentices. The skill group U comprises ND and HS, M comprises VT and HSVT; H comprises TC and UD. The 50th and the 80th wage percentile for H cannot be reported because the wage data is right censored.

Table 4.13: **Wage Inequality for Men Between Skill Groups for 1984 and 1997**

Year	Groups	At percentile	Orig. data	IP1	IP2A	IP2B	IP3
1984	M-U	50	0.126	0.127	0.117	0.117	0.117
	M-U	20	0.118	0.095	0.095	0.095	0.095
	H-M	20	0.463	0.457	0.471	0.471	0.464
1997	M-U	50	0.156	0.140	0.173	0.152	0.139
	M-U	20	0.163	0.187	0.219	0.185	0.168
	H-M	20	0.438	0.423	0.450	0.446	0.428
Change	M-U	50	0.030	0.013	0.056	0.035	0.022
84/97	M-U	20	0.045	0.092	0.124	0.090	0.073
	H-M	20	-0.026	-0.034	-0.022	-0.026	-0.036

Notes: The table contains differences in log daily wages between skill groups at specific wage percentiles based on the wage values from table 4.12 .

Table 4.14: **Wage Inequality for Men Within Skill Groups for 1984 and 1997**

Year	Skill group	Measure	Orig. data	IP1	IP2A	IP2B	IP3
1984	U	50%-20%	0.193	0.182	0.193	0.193	0.193
		80%-50%	0.179	0.181	0.179	0.179	0.187
	M	50%-20%	0.201	0.214	0.214	0.214	0.214
		80%-50%	0.276	0.264	0.271	0.271	0.264
1997	U	50%-20%	0.228	0.269	0.264	0.260	0.257
		80%-50%	0.197	0.194	0.203	0.200	0.205
	M	50%-20%	0.222	0.222	0.217	0.226	0.228
		80%-50%	0.291	0.280	0.280	0.277	0.278
Change	U	50%-20%	0.036	0.086	0.071	0.067	0.065
84/97		80%-50%	0.019	0.014	0.024	0.021	0.017
	M	50%-20%	0.021	0.008	0.003	0.012	0.014
		80%-50%	0.014	0.015	0.009	0.005	0.014

Notes: The table contains differences in log daily wages within skill groups between the respective percentiles based on the wage values from table 4.12.

Table 4.15: **Mincer-type Earnings Regression (Tobit) 1**

	Original data		IP1		IP2A		IP2B		IP3	
ND	-0.205	(0.002)	-0.193	(0.003)	-0.196	(0.003)	-0.197	(0.003)	-0.187	(0.003)
HS	0.057	(0.018)	0.007	(0.021)	0.016	(0.019)	0.016	(0.019)	0.035	(0.020)
HSVT	0.196	(0.007)	0.123	(0.005)	0.176	(0.006)	0.174	(0.006)	0.135	(0.006)
TC	0.430	(0.005)	0.414	(0.005)	0.433	(0.005)	0.434	(0.005)	0.424	(0.005)
UD	0.473	(0.005)	0.479	(0.005)	0.496	(0.005)	0.497	(0.005)	0.485	(0.005)
age/10	0.211	(0.005)	0.193	(0.005)	0.192	(0.005)	0.194	(0.005)	0.189	(0.005)
age_sq/100	-0.016	(0.001)	-0.014	(0.001)	-0.014	(0.001)	-0.014	(0.001)	-0.014	(0.001)
intercept	4.554	(0.009)	4.557	(0.009)	4.570	(0.009)	4.562	(0.009)	4.563	(0.009)
lnsigma	-1.144	(0.004)	-1.100	(0.004)	-1.115	(0.004)	-1.109	(0.004)	-1.101	(0.004)
N	145483		157935		156042		157668		156757	
censored	16789		17549		17466		17537		17407	

Notes: Dependent variable log daily wage, which is right censored at the social security threshold. Men in West Germany working fulltime 1995, no apprentices. The omitted education is VT. Spells weighted with their length. Robust standard errors clustered at the person level are in parentheses.

Table 4.16: **Mincer-type Earnings Regression (Tobit) 2**

	Original data		IP1		IP2A		IP2B		IP3	
ND	-0.125	(0.002)	-0.112	(0.003)	-0.122	(0.003)	-0.118	(0.003)	-0.109	(0.003)
HS	-0.018	(0.017)	-0.054	(0.020)	-0.057	(0.018)	-0.052	(0.018)	-0.035	(0.019)
HSVT	0.086	(0.006)	0.035	(0.005)	0.072	(0.006)	0.070	(0.006)	0.042	(0.005)
TC	0.245	(0.005)	0.221	(0.005)	0.240	(0.005)	0.240	(0.005)	0.230	(0.005)
UD	0.320	(0.005)	0.307	(0.005)	0.331	(0.005)	0.330	(0.005)	0.314	(0.005)
age/10	0.174	(0.005)	0.165	(0.005)	0.163	(0.005)	0.165	(0.005)	0.161	(0.005)
age_sq/100	-0.013	(0.001)	-0.012	(0.001)	-0.012	(0.001)	-0.012	(0.001)	-0.011	(0.001)
foreign	-0.064	(0.003)	-0.081	(0.003)	-0.071	(0.003)	-0.075	(0.003)	-0.080	(0.003)
farmer	-0.212	(0.010)	-0.225	(0.010)	-0.222	(0.010)	-0.223	(0.009)	-0.222	(0.010)
service worker	-0.032	(0.006)	-0.056	(0.006)	-0.051	(0.006)	-0.055	(0.006)	-0.056	(0.006)
sales worker	0.171	(0.006)	0.176	(0.006)	0.172	(0.005)	0.172	(0.005)	0.176	(0.006)
clerical worker	0.227	(0.003)	0.244	(0.003)	0.236	(0.003)	0.237	(0.003)	0.244	(0.003)
admin worker	0.281	(0.003)	0.297	(0.003)	0.287	(0.003)	0.289	(0.003)	0.296	(0.003)
agriculture	-0.012	(0.005)	-0.012	(0.005)	-0.012	(0.005)	-0.013	(0.005)	-0.012	(0.005)
basic industry	0.011	(0.003)	0.008	(0.003)	0.010	(0.003)	0.010	(0.003)	0.008	(0.003)
consumer goods	-0.090	(0.003)	-0.093	(0.003)	-0.090	(0.003)	-0.091	(0.003)	-0.093	(0.003)
food industry	-0.112	(0.005)	-0.117	(0.005)	-0.115	(0.005)	-0.116	(0.005)	-0.117	(0.005)
main construction	-0.047	(0.003)	-0.052	(0.003)	-0.050	(0.003)	-0.051	(0.003)	-0.053	(0.003)
constr completion	-0.128	(0.004)	-0.126	(0.004)	-0.126	(0.004)	-0.127	(0.004)	-0.125	(0.004)
trade	-0.179	(0.004)	-0.190	(0.003)	-0.186	(0.003)	-0.188	(0.003)	-0.190	(0.003)
transport & comm	-0.131	(0.004)	-0.152	(0.004)	-0.149	(0.004)	-0.150	(0.004)	-0.153	(0.004)
business services	-0.113	(0.004)	-0.127	(0.004)	-0.126	(0.004)	-0.128	(0.004)	-0.129	(0.004)
consumer services	-0.337	(0.009)	-0.375	(0.009)	-0.365	(0.009)	-0.374	(0.009)	-0.376	(0.009)
education	-0.195	(0.004)	-0.202	(0.004)	-0.202	(0.004)	-0.203	(0.004)	-0.204	(0.004)
public admin	-0.183	(0.004)	-0.186	(0.004)	-0.185	(0.004)	-0.185	(0.004)	-0.188	(0.004)
intercept	4.653	(0.009)	4.659	(0.009)	4.667	(0.009)	4.662	(0.009)	4.665	(0.009)
lnsigma	-1.239	(0.004)	-1.207	(0.004)	-1.216	(0.004)	-1.212	(0.004)	-1.209	(0.004)
N	141,860		153,431		151,769		153,119		152,258	
censored	16,602		17,302		17,228		17,294		17,162	

Notes: Dependent variable log daily wage, which is right censored at the social security threshold. Men in West Germany working fulltime 1995, no apprentices. The omitted education is VT, omitted occupation salaried employee and omitted industry investment goods industry. Spells weighted with their length. Robust standard errors clustered at the person level are in parenthesis.

Table 4.17: **Earnings and Misreports – Mincer-type Earnings Regression (Tobit)**

	IP1		IP2A		IP2B		IP3	
ND	-0.120	(0.003)	-0.120	(0.003)	-0.119	(0.003)	-0.117	(0.003)
HS	-0.032	(0.020)	-0.039	(0.018)	-0.034	(0.018)	-0.017	(0.019)
HSVT	0.085	(0.005)	0.103	(0.006)	0.101	(0.006)	0.086	(0.005)
TC	0.244	(0.005)	0.253	(0.005)	0.252	(0.005)	0.250	(0.005)
UD	0.324	(0.005)	0.340	(0.005)	0.339	(0.005)	0.328	(0.005)
age/10	0.171	(0.005)	0.165	(0.005)	0.169	(0.005)	0.169	(0.005)
age_sq/100	-0.012	(0.001)	-0.012	(0.001)	-0.012	(0.001)	-0.012	(0.001)
reportmiss	-0.120	(0.004)	-0.106	(0.005)	-0.112	(0.004)	-0.120	(0.004)
underreport	-0.110	(0.003)	-0.106	(0.004)	-0.104	(0.004)	-0.116	(0.003)
overreport			-0.181	(0.012)	-0.211	(0.017)	0.008	(0.020)
foreign	-0.068	(0.003)	-0.065	(0.003)	-0.067	(0.003)	-0.068	(0.003)
farmer	-0.217	(0.009)	-0.216	(0.009)	-0.217	(0.009)	-0.215	(0.009)
service worker	-0.047	(0.006)	-0.044	(0.006)	-0.047	(0.006)	-0.046	(0.006)
sales worker	0.165	(0.005)	0.165	(0.005)	0.165	(0.005)	0.165	(0.005)
clerical worker	0.229	(0.003)	0.227	(0.003)	0.227	(0.003)	0.230	(0.003)
admin worker	0.282	(0.003)	0.279	(0.003)	0.280	(0.003)	0.282	(0.003)
agriculture	-0.014	(0.005)	-0.013	(0.005)	-0.014	(0.005)	-0.013	(0.005)
basic industry	0.012	(0.003)	0.013	(0.003)	0.013	(0.003)	0.013	(0.003)
consumer goods	-0.089	(0.003)	-0.087	(0.003)	-0.088	(0.003)	-0.088	(0.003)
food industry	-0.112	(0.005)	-0.111	(0.005)	-0.111	(0.005)	-0.112	(0.005)
main construction	-0.039	(0.003)	-0.039	(0.003)	-0.038	(0.003)	-0.039	(0.003)
constr completion	-0.125	(0.004)	-0.124	(0.004)	-0.124	(0.004)	-0.123	(0.004)
trade	-0.178	(0.003)	-0.177	(0.003)	-0.177	(0.003)	-0.178	(0.003)
transport & comm	-0.140	(0.004)	-0.138	(0.004)	-0.138	(0.004)	-0.139	(0.004)
business services	-0.118	(0.004)	-0.117	(0.004)	-0.118	(0.004)	-0.119	(0.004)
consumer services	-0.359	(0.009)	-0.350	(0.009)	-0.357	(0.009)	-0.359	(0.009)
education	-0.198	(0.004)	-0.198	(0.004)	-0.199	(0.004)	-0.198	(0.004)
public admin	-0.182	(0.004)	-0.182	(0.004)	-0.182	(0.004)	-0.183	(0.004)
intercept	4.660	(0.009)	4.664	(0.009)	4.664	(0.009)		
lnsigma	-1.217	(0.004)	-1.225	(0.004)	-1.220	(0.004)	-1.218	(0.004)
N	153,431		151,769		153,199		152,258	
censored	17,302		17,228		17,294		17,162	

Notes: Dependent variable: log daily wage, which is right censored at the social security threshold. Men in West Germany working fulltime 1995, no apprentices. The omitted education is VT, omitted occupation salaried employee and omitted industry investment goods industry. Spells weighted with their length. Robust standard errors clustered at the person level are in parentheses. *reportmiss, underreport* and *overreport* are defined in comparison to the original data. No *overreport* for IP1 since this procedure assumes there are no overreports for persons over 17 years.

Table 4.18: **Probit Regression of Underreport Compared to IP2A**

Regressors	Marg eff	Robust SE	Regressors	Marg eff	Robust SE
≤19 years	-0.048	(0.000)**	1-2 reports by empl	0.037	(0.001)**
30-39 years	0.015	(0.001)**	3-5 reports by empl	0.020	(0.001)**
40-49 years	0.007	(0.001)**	spell≤30 days	0.016	(0.001)**
50-59 years	-0.005	(0.001)**	30<spell≤180 days	0.011	(0.000)**
60+ years	-0.017	(0.001)**	year 75	-0.037	(0.000)**
female	-0.007	(0.001)**	year 76	-0.034	(0.000)**
married	-0.003	(0.000)**	year 77	-0.029	(0.000)**
foreign	-0.020	(0.001)**	year 78	-0.022	(0.000)**
trainee	0.029	(0.001)**	year 79	-0.016	(0.000)**
non-skilled worker	0.057	(0.001)**	year 80	-0.012	(0.000)**
skilled worker	-0.028	(0.001)**	year 81	-0.008	(0.000)**
master craftsman/foreman	-0.024	(0.001)**	year 82	-0.006	(0.000)**
home worker	0.042	(0.008)**	year 83	-0.005	(0.000)**
part time ≤18h	0.023	(0.003)**	year 84	-0.002	(0.000)**
part time >18h	0.012	(0.001)**	year 86	0.001	(0.000)*
farmers/farm managers	0.012	(0.002)**	year 87	0.002	(0.000)**
service workers	-0.002	(0.001)*	year 88	0.003	(0.000)**
sales workers	-0.007	(0.001)**	year 89	0.004	(0.001)**
clerical workers	-0.002	(0.001)	year 90	0.005	(0.001)**
admin/profes/techn staff	-0.000	(0.001)	year 91	0.005	(0.001)**
agriculture	-0.006	(0.002)**	year 92	0.005	(0.001)**
basic industry	-0.002	(0.001)	year 93	0.005	(0.001)**
consumer goods industry	0.006	(0.001)**	year 94	0.004	(0.001)**
food industry	-0.002	(0.001)	year 95	0.003	(0.001)**
main construction trade	0.026	(0.002)**	year 96	0.001	(0.001)
construction completion trade	0.002	(0.002)	year 97	-0.001	(0.001)
trade	0.006	(0.001)**			
transport and communication	-0.003	(0.001)*			
business services	0.007	(0.001)**			
consumer services	0.003	(0.001)			
education, non profit org	0.002	(0.001)			
public administration	-0.002	(0.001)			
observed Prob	0.060				
predicted Prob at $\bar{x}$	0.047				
N	6,352,330				
Pseudo R^2	0.078				

Notes: Dependent variable: dummy for reported education lower than imputed education (IP2A). Estimation based on all employment spells in West Germany. Base category: 20-29 years, male, not married, German, working fulltime as a salaried employee, occupation group operatives/craft, investment goods industry, more than five reports by the employer about the employee, spell longer than 180 days, 1985. Intercept included in estimation. Robust standard errors with clustering at the person level. * significant at 5%, ** significant at 1%.

Appendix A

Additional Appendix to Chapter 1

A.1 Estimation Results for the Propensity Score

A.1.1 Sample Sizes and Variable Definitions

Sample Sizes

Cohort 86/87			
	Stratum 1	Stratum 2	Stratum 3
Waiting	20,153	9,440	6,364
PF	74	60	69
SPST	503	257	176
RT	172	101	71

Cohort 93/94			
	Stratum 1	Stratum 2	Stratum 3
Waiting	24,223	13,751	9,244
PF	102	102	86
SPST	528	481	669
RT	198	138	106

Variable Definitions

Table A.1: Variable Definitions

Label	Definition
Personal Attributes	
aXXYY	Age at start of unemployment $\geq$XX and $\leq$ YY
age	Age at start of unemployment
female	Female
foreign	No German citizenship
kids	Has dependent children
married	Married
qual_u	No vocational training degree
qual_l	No vocational training degree or education information missing
qual_m	Vocational training degree
qual_h	University/College degree
Last Employment	
BER1	Apprentice
BER2	Blue Collar Worker
BER3	White Collar Worker
BER4	Worker at home with low hours or BER missing
BER5	Part–time working
pearn	Daily earnings $\geq$ 15 Euro per day in 1995 Euro
earncens	Earnings censored at social security taxation threshold
earn	Daily earnings if pearn=1 and earncens=0, otherwise zero
logearn	log(earn) if pearn=1 and earncens=0, otherwise zero
logearnsq	logearn squared
earnp90	Daily earnings above 90th percentile
Last Employer	
industry1	Agriculture
industry2	Basic materials
industry3	Metal, vehicles, electronics
industry4	Light industry
industry5	Construction
industry6	Production oriented services, trade, banking

<continued on next page>

Table A.1: Variable Definitions <continued>

Label	Definition
industry7	Consumer oriented services, organization and social services
frmsize1	Firm Size (employment) missing or ≤ 10
frmsize2	Firm Size (employment) > 10 and ≤ 200
frmsize3	Firm Size (employment) > 200 and ≤ 500
frmsize4	Firm Size (employment) > 500
Employment and Program History	
preexM	Employed M (M=6, 12, 24) month before unemployment starts
preex60cumst	Number of months employed in the last 60 months before unemployment starts, standardized
preex60sq	preex60cumst squared
pretxY	Participation in any ALMP program reported in our data in year(s) Y (Y=1, 2, 3–5) before unemployment starts
Regional Information	
state6	Schleswig-Holstein/Hamburg
state7	Niedersachsen-Bremen
state8	Nordrhein-Westfalen
state9	Hessen
state10	Rheinland-Pfalz/ Saarland
state11	Baden-Württemberg
state12	Bayern
denst	population density (standardized)
densq	denst squared
R1	Population density <100 inhabitants per square kilometer, Rural area
R2	Population density $\geq$100 and $<$ 150, Medium population density
R3	Population density $\geq$150 and $<$ 400, Dense area
R4	Population density $\geq$400, Metropolitan area
ur	Unemployment rate at district level (Kreis), 80s
ursq	ur squared
urtb	Unemployment rate at district level (Kreis), 90s

<continued on next page>

Table A.1: Variable Definitions <continued>

Label	Definition
urtbsq	urtb squared
urtb100	urtb/100
Calendar Time of Entry into Unemployment	
tnull	First unemployment month (months counted from January 1960)
uentry	First unemployment month (months counted from January 1986 (1993) in the 80s (90s))
uentry2	uentry squared
yYY	Unemployment begins in year YY
qQ	Unemployment begins in quarter Q of the year
yYYqQ	Unemployment begins in quarter Q of year YY
Interaction of Variables	
f_	female
for_	foreign

All variables are defined at the time of entry into unemployment and constant during the unemployment spell.

A.1.2 Results of Propensity Score Estimations and Balancing Tests

Remark: The propensity score tables show the estimated coefficients of the probit regressions of the conditional probability to participate in the first of the two treatments mentioned in the header. The estimations are carried out separately for each time window of elapsed unemployment duration (Stratum 1, 2, and 3). Standard errors are in parentheses. *, **, *** means significant at the 10%, 5%, 1% level, respectively, in a two–sided test. Each probit table is followed by a table indicating how many regressors pass the Smith/Todd (2005) balancing test at different significance levels using a cubic and a quartic of the propensity score, respectively. Graphs with the densities of the propensity scores are in the next subsection.

Propensity Score Estimates, Treatment PF vs Waiting, Cohort 86/87

	Stratum 1	Stratum 2	Stratum 3
state10	0.327 (0.140)**		
state79	0.479 (0.090)***		
a2529			0.362 (0.315)
a2534		0.104 (0.197)	
a3034	-0.073 (0.124)		0.829 (0.307)***
a3539	-0.020 (0.136)	0.050 (0.233)	
a3544			0.734 (0.333)**
a4044	-0.129 (0.154)	0.107 (0.239)	
a4549	-0.070 (0.141)	0.240 (0.223)	0.487 (0.267)*
a5055	-0.308 (0.177)*		
ur	0.094 (0.067)	0.149 (0.085)*	0.265 (0.095)***
ursq	-0.005 (0.003)*	-0.007 (0.004)*	-0.011 (0.004)***
densq			0.094 (0.042)**
denst			-0.183 (0.081)**
earn			-0.021 (0.009)**
f_BER3		0.191 (0.198)	
f_a3034		0.546 (0.218)**	
f_a3539		0.439 (0.306)	
f_a3544	0.355 (0.200)*		
f_preex60cumst			-0.089 (0.090)
female	-0.388 (0.126)***	-0.709 (0.202)***	-0.468 (0.117)***
frmsize23	0.211 (0.089)**	0.148 (0.116)	
frmsize4		0.273 (0.154)*	
logearn	0.018 (0.071)	0.038 (0.066)	
logearnsq			0.099 (0.048)**
married			-0.218 (0.107)**
pearn	-0.374 (0.384)		
preex12			-0.305 (0.130)**
preex24	0.205 (0.117)*	-0.181 (0.098)*	
preex60cumst	-0.063 (0.057)		0.189 (0.073)***
preex60sq	0.084 (0.037)**		0.113 (0.043)***
pretx1			-0.423 (0.281)
pretx2			0.652 (0.208)***
pretx35			-0.232 (0.206)
qual_l		-0.528 (0.334)	
qual_l_a2539		0.866 (0.356)**	
qual_m_a3544			0.088 (0.189)
qual_m_a4555			0.036 (0.263)
uentry		0.036 (0.028)	0.010 (0.007)
uentry2		-0.002 (0.001)	
y86q2	0.705 (0.317)**		

Propensity Score Estimates, PF vs Waiting, Cohort 86/87 – continued			
	Stratum 1	Stratum 2	Stratum 3
y86q34	0.779 (0.290)***		
y87q1	0.812 (0.292)***		
y87q2	0.926 (0.303)***		
y87q3	1.050 (0.295)***		
y87q4	0.848 (0.296)***		
_cons	-3.935 (0.535)***	-3.496 (0.555)***	-4.605 (0.677)***
N	20227	9500	6433

Balancing Tests, Treatment PF vs Waiting, Cohort 86/87				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	21	22	23	23
Stratum 2	16	17	17	18
Stratum 3	16	21	21	22
Quartic of Pscore				
Stratum 1	15	17	17	23
Stratum 2	15	15	16	18
Stratum 3	12	14	17	22

Propensity Score Estimates, Treatment SPST vs Waiting, Cohort 86/87			
	Stratum 1	Stratum 2	Stratum 3
BER1	-0.003 (0.194)		
BER2		-0.067 (0.126)	-0.117 (0.121)
BER3	0.302 (0.054)***	0.275 (0.170)	0.202 (0.112)*
BER3_a2539		-0.120 (0.127)	
industry3	0.317 (0.065)***		0.163 (0.146)
industry4	0.098 (0.080)		0.081 (0.160)
industry5	-0.161 (0.076)**		
industry6	0.230 (0.050)***		0.112 (0.129)
industry7			0.014 (0.132)
a2529	0.404 (0.091)***		1.075 (0.235)***
a3034	0.382 (0.094)***	-0.030 (0.102)	1.192 (0.236)***
a3539	0.445 (0.096)***	0.110 (0.226)	
a3544			0.814 (0.256)***
a4044	0.213 (0.118)*	-0.096 (0.231)	
a4549	0.233 (0.103)**	-0.023 (0.238)	0.727 (0.191)***
a5055		-0.504 (0.248)**	

Propensity Score Estimates, SPST vs Waiting, Cohort 86/87 – continued			
	Stratum 1	Stratum 2	Stratum 3
ur			0.134 (0.059)**
ursq			-0.006 (0.003)**
denst			0.038 (0.034)
earncens	0.372 (0.232)	0.624 (0.240)***	
f_BER2	-0.292 (0.094)***		
f_BER3		0.063 (0.134)	
f_industry7		-0.224 (0.095)**	
f_a2529		-0.069 (0.126)	
f_a3544			0.101 (0.168)
f_a4044	0.355 (0.125)***		
f_a4555			0.626 (0.205)***
female	0.105 (0.056)*	0.012 (0.113)	-0.217 (0.100)**
for_age		-0.009 (0.003)***	
foreign	-0.109 (0.083)		-0.327 (0.142)**
logearn	0.107 (0.044)**		
logearnsq		0.032 (0.010)***	
m_industry5			-0.755 (0.293)***
married	-0.120 (0.042)***	-0.191 (0.059)***	-0.238 (0.073)***
preex12	0.129 (0.049)***		
preex60cumst	-0.045 (0.022)**	-0.033 (0.029)	
preex60sq			0.002 (0.033)
preex60sq_a3544		0.125 (0.039)***	0.145 (0.056)***
pretx1	0.204 (0.096)**	0.235 (0.122)*	0.372 (0.150)**
pretx2	0.071 (0.095)		
pretx35	0.066 (0.072)		
qual_h	0.277 (0.108)**	0.387 (0.196)**	
qual_h_a3544		-0.224 (0.299)	
qual_h_a4555		0.252 (0.349)	
qual_m	0.261 (0.072)***	0.316 (0.143)**	
qual_m_a3544		-0.224 (0.208)	
qual_m_a4555		0.077 (0.240)	
uentry	0.024 (0.010)**		
uentry2	-0.001 (0.000)*	-0 (0.000)***	
y86q2			-0.267 (0.141)*
y86q3			-0.046 (0.127)
y86q4			-0.120 (0.131)
y87q1			-0.245 (0.134)*
y87q2			-0.188 (0.134)
y87q3			-0.399 (0.146)***
y87q4			-0.016 (0.124)
_cons	-3.345 (0.213)***	-2.508 (0.233)***	-3.316 (0.419)***

Propensity Score Estimates, SPST vs Waiting, Cohort 86/87 – continued			
	Stratum 1	Stratum 2	Stratum 3
N	20656	9697	6540

Balancing Tests, Treatment SPST vs Waiting, Cohort 86/87				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	24	26	26	27
Stratum 2	20	25	26	26
Stratum 3	24	27	27	29
Quartic of Pscore				
Stratum 1	19	23	25	27
Stratum 2	19	21	24	26
Stratum 3	20	22	23	29

Propensity Score Estimates, Treatment RT vs Waiting, Cohort 86/87

	Stratum 1	Stratum 2	Stratum 3
BER3			-0.099 (0.121)
state10	-0.074 (0.128)	-0.275 (0.186)	
state11	0.081 (0.101)	-0.088 (0.140)	
state12		-0.245 (0.149)*	
state6		0.026 (0.132)	-0.030 (0.183)
state612	-0.175 (0.086)**		
state7	-0.089 (0.099)	-0.141 (0.128)	
state710			0.347 (0.108)***
state9	0.200 (0.100)**	0.077 (0.139)	
a2529	0.837 (0.146)***	0.994 (0.301)***	
a2534			0.760 (0.356)**
a3034	0.848 (0.150)***	1.062 (0.303)***	
a3539		0.666 (0.320)**	0.669 (0.425)
a3544	0.658 (0.151)***		
a4044		0.682 (0.326)**	0.400 (0.431)
a4549		0.449 (0.339)	
densq	-0.036 (0.029)	-0.170 (0.077)**	
denst	0.110 (0.049)**	0.119 (0.072)*	0.061 (0.047)
f_densq		0.072 (0.080)	
f_preex60cumst		-0.137 (0.076)*	
f_qual_h	0.341 (0.144)**		
f_uentry			0.009 (0.014)

Propensity Score Estimates, RT vs Waiting, Cohort 86/87 – continued			
	Stratum 1	Stratum 2	Stratum 3
female	-0.112 (0.070)	-0.585 (0.219)***	-0.196 (0.196)
foreign	-0.340 (0.130)***	-0.577 (0.205)***	-0.209 (0.188)
logearn	0.074 (0.061)	0.061 (0.055)	
logearnsq			0.028 (0.017)*
m_BER2		-0.420 (0.207)**	
m_BER3		-0.318 (0.228)	
pearn	-0.290 (0.329)		-0.633 (0.355)*
preex12	0.147 (0.073)**		
preex60cumst	-0.150 (0.035)***	-0.114 (0.064)*	
preex60cumst_a2534			-0.072 (0.055)
preex60cumst_a3544	0.164 (0.066)**	0.233 (0.104)**	-0.018 (0.079)
preex60sq		-0.061 (0.039)	
pretx1	0.267 (0.118)**		
qual_h			0.156 (0.263)
qual_h_a3544			0.587 (0.422)
qual_m			-0.138 (0.453)
qual_m_a2534			0.316 (0.472)
qual_m_a3544			0.423 (0.526)
uentry		-0.003 (0.006)	0.001 (0.010)
y86q2			0.362 (0.147)**
y86q23	0.155 (0.120)		
y86q4	0.208 (0.125)*		
y87q1	0.192 (0.119)		
y87q2	0.235 (0.136)*		
y87q3	0.371 (0.124)***		
y87q4	0.342 (0.119)***		
_cons	-3.308 (0.292)***	-2.609 (0.411)***	-2.950 (0.441)***
N	20325	9541	6435

Balancing Tests, Treatment RT vs Waiting, Cohort 86/87				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	24	24	24	25
Stratum 2	19	22	23	24
Stratum 3	18	20	20	21
Quartic of Pscore				
Stratum 1	15	19	22	25
Stratum 2	11	12	15	24
Stratum 3	13	14	17	21

Propensity Score Estimates, Treatment PF vs SPST, Cohort 86/87

	Stratum 1	Stratum 2	Stratum 3
R1	0.528 (0.226)**		
a2534	-0.197 (0.147)		
densq			0.134 (0.078)*
denst			-0.304 (0.139)**
f_BER3	-0.621 (0.275)**		
f_preex6		-1.006 (0.390)***	
f_preex60sq		-0.261 (0.129)**	
f_tnull	0.037 (0.011)***		
female		1.510 (0.741)**	-0.868 (0.200)***
foreign	0.437 (0.256)*	0.691 (0.318)**	
logearn	0.729 (0.330)**		1.170 (0.446)***
logearnsq	-0.181 (0.069)***	-0.066 (0.034)**	-0.232 (0.086)***
m_logearn		0.301 (0.189)	
m_tnull	0.039 (0.012)***		
preex6	-0.342 (0.167)**		
qual_mh		-0.753 (0.235)***	
qual_u			0.651 (0.227)***
tnull		0.014 (0.013)	
_cons	-13.225 (3.797)***	-4.685 (4.442)	-1.449 (0.647)**
N	577	317	245

Balancing Tests, Treatment PF vs SPST, Cohort 86/87

	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	8	9	9	9
Stratum 2	8	8	8	8
Stratum 3	5	5	6	6
Quartic of Pscore				
Stratum 1	8	9	9	9
Stratum 2	8	8	8	8
Stratum 3	6	6	6	6

Propensity Score Estimates, Treatment PF vs RT, Cohort 86/87

	Stratum 1	Stratum 2	Stratum 3
state10	0.831 (0.335)**		
state12	0.580 (0.299)*		
state7	0.868 (0.266)***		
state9	0.634 (0.270)**		
industry7		-0.707 (0.275)**	
a2529		-1.428 (0.418)***	-1.199 (0.326)***
a3034		-1.033 (0.410)**	-0.616 (0.318)*
a3539		-0.592 (0.473)	-0.663 (0.349)*
a4044		-0.822 (0.543)	
ur100			30.948 (21.463)
ursq			-141.761 (95.828)
densq	0.103 (0.099)	0.352 (0.147)**	
denst	-0.221 (0.157)	-0.524 (0.187)***	
f_a2534	-1.197 (0.386)***		
f_preex12	-0.317 (0.489)		
f_preex24	0.924 (0.546)*		
female	-1.041 (0.705)		-0.544 (0.239)**
foreign		1.455 (0.437)***	
m_a2534	-1.121 (0.376)***		
m_a3544	-0.868 (0.403)**		
married		-0.356 (0.237)	
tnull	0.007 (0.015)		
_cons	-2.379 (4.870)	0.632 (0.411)	-0.713 (1.168)
N	246	161	140

Balancing Tests, Treatment PF vs RT, Cohort 86/87				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	12	12	13	13
Stratum 2	8	8	9	9
Stratum 3	6	6	6	6
Quartic of Pscore				
Stratum 1	12	13	13	13
Stratum 2	8	8	9	9
Stratum 3	6	6	6	6

Propensity Score Estimates, Treatment RT vs SPST, Cohort 86/87

	Stratum 1	Stratum 2	Stratum 3
BER2			0.327 (0.175)*
BER3	-0.474 (0.118)***		
state1012		-0.450 (0.209)**	
a2529			1.051 (0.382)***
a2534	1.004 (0.251)***	1.033 (0.263)***	
a3034			0.914 (0.393)**
a3539			1.132 (0.406)***
a3544	0.701 (0.265)***	0.392 (0.293)	
a4044			0.759 (0.451)*
densq	-0.111 (0.053)**	-0.119 (0.068)*	
denst	0.210 (0.086)**		
earnp90		-0.411 (0.265)	
f_preex60cumst		-0.465 (0.162)***	
f_preex60sq		-0.269 (0.093)***	
preex60cumst	-0.170 (0.070)**		
preex60sq	-0.031 (0.053)		
qual_h		-0.776 (0.330)**	
qual_m		-0.710 (0.229)***	
qual_u	0.463 (0.185)**		
y87	0.299 (0.112)***		
_cons	-1.453 (0.267)***	-0.372 (0.316)	-1.624 (0.359)***
N	675	358	247

Balancing Tests, Treatment RT vs SPST, Cohort 86/87

	P-values>.1	P-values>.05	P-values>.01	Regressors
	Cubic of Pscore			
Stratum 1	9	9	9	9
Stratum 2	8	9	9	9
Stratum 3	5	5	5	5
	Quartic of Pscore			
Stratum 1	9	9	9	9
Stratum 2	8	9	9	9
Stratum 3	5	5	5	5

Propensity Score Estimates, Treatment PF vs Waiting, Cohort 93/94

	Stratum 1	Stratum 2	Stratum 3
BER2	0.225 (0.154)	0.192 (0.152)	-0.157 (0.197)
BER3	0.275 (0.153)*	0.211 (0.156)	-0.492 (0.241)**
state10	0.216 (0.127)*		0.078 (0.186)
state1112			0.068 (0.143)
state67			0.367 (0.120)***
state7	0.263 (0.101)***	0.259 (0.091)***	
state9	0.207 (0.108)*		0.085 (0.181)
industry6			-0.120 (0.126)
a3034	0.130 (0.104)	-0.276 (0.122)**	
a3539	0.124 (0.114)	0.046 (0.110)	0.373 (0.129)***
a4044		-0.075 (0.132)	0.368 (0.134)***
a4049	0.274 (0.097)***		
a4549		-0.086 (0.139)	0.311 (0.147)**
a5055	-0.304 (0.173)*	-0.276 (0.138)**	0.071 (0.163)
urtb	-0.015 (0.016)		0.275 (0.130)**
urtbsq			-0.013 (0.007)**
densq		0.012 (0.034)	0.087 (0.040)**
denst	-0.013 (0.038)	-0.074 (0.060)	-0.211 (0.080)***
f_BER3			1.107 (0.260)***
f_industry6			0.338 (0.204)*
f_a4055		0.236 (0.156)	
f_a5055			0.518 (0.223)**
f_logearn	0.141 (0.128)	0.109 (0.111)	
f_qual_m	0.670 (0.221)***		
f_uentry	0.018 (0.010)*		
female	-1.233 (0.559)**	0.014 (0.399)	-0.977 (0.219)***
foreign		0.070 (0.100)	-0.202 (0.138)
logearnsq	-0.009 (0.010)		0.004 (0.011)
m_logearn		0.152 (0.089)*	
m_pretx35			0.362 (0.137)***
married			-0.215 (0.097)**
pearn		-0.794 (0.420)*	
preex12			-0.171 (0.117)
preex24	0.119 (0.079)		
preex60cumst		-0.014 (0.041)	0.119 (0.058)**
preex60sq	-0.028 (0.046)		
pretx35	0.240 (0.091)***		
qual_m		0.205 (0.093)**	
uentry	-0.007 (0.007)		-0.012 (0.007)*
y93q2		-0.251 (0.148)*	
y93q3		-0.065 (0.127)	

Propensity Score Estimates, PF vs Waiting, Cohort 93/94 – continued			
	Stratum 1	Stratum 2	Stratum 3
y93q4		-0.162 (0.132)	
y94q1		-0.117 (0.132)	
y94q2		-0.175 (0.150)	
y94q3		-0.125 (0.139)	
y94q4		-0.319 (0.155)**	
_cons	-2.824 (0.264)***	-2.391 (0.380)***	-3.505 (0.686)***
N	24325	13853	9330

Balancing Tests, Treatment PF vs Waiting, Cohort 93/94				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	19	19	19	20
Stratum 2	23	24	25	25
Stratum 3	23	25	25	26
Quartic of Pscore				
Stratum 1	14	18	19	20
Stratum 2	17	17	19	25
Stratum 3	20	22	24	26

Propensity Score Estimates, Treatment SPST vs Waiting, Cohort 93/94			
	Stratum 1	Stratum 2	Stratum 3
BER1			0.481 (0.169)***
BER2	-0.124 (0.073)*	-0.088 (0.082)	0.039 (0.077)
BER3	0.270 (0.071)***	0.110 (0.083)	0.151 (0.080)*
state10	0.199 (0.070)***	0.222 (0.085)***	
state11	-0.123 (0.063)*	0.031 (0.069)	
state12	-0.027 (0.055)	0.096 (0.069)	
state6	-0.047 (0.075)	-0.003 (0.088)	
state7	-0.100 (0.065)	0.021 (0.073)	
state9	-0.118 (0.073)	-0.128 (0.085)	
industry3	0.147 (0.076)*	-0.058 (0.086)	0.038 (0.080)
industry4	-0.010 (0.090)	0.076 (0.094)	0.007 (0.091)
industry5	-0.057 (0.094)	-0.377 (0.120)***	-0.223 (0.106)**
industry6	0.073 (0.072)	-0.003 (0.080)	0.026 (0.076)
industry7	-0.127 (0.079)	-0.195 (0.088)**	-0.085 (0.081)
a3034	0.019 (0.054)	0.097 (0.063)	0.155 (0.062)**
a3539	-0.065 (0.069)	0.162 (0.068)**	0.189 (0.075)**

Propensity Score Estimates, SPST vs Waiting, Cohort 93/94 – continued			
	Stratum 1	Stratum 2	Stratum 3
a4044		-0.483 (0.157)***	-0.283 (0.119)**
a4049	-0.085 (0.064)		
a4549		-0.654 (0.162)***	-0.482 (0.128)***
a5055	-0.460 (0.082)***	-0.891 (0.161)***	-0.914 (0.128)***
urtb			-0.014 (0.008)*
densq		0.036 (0.018)**	
denst		-0.033 (0.037)	
earncens	0.023 (0.190)	0.362 (0.335)	-0.298 (0.306)
f_industry5		0.851 (0.235)***	
f_a2534	-0.192 (0.078)**		-0.402 (0.107)***
f_a3544			-0.185 (0.109)*
f_for_a2539		-0.181 (0.093)*	
f_married			-0.183 (0.066)***
f_qual_h			-0.623 (0.177)***
f_qual_m			-0.249 (0.086)***
female	0.034 (0.058)	0.009 (0.081)	0.428 (0.118)***
for_a2534	-0.301 (0.098)***		
for_a2539		-0.205 (0.084)**	
for_a3544	-0.209 (0.119)*		
foreign			-0.331 (0.062)***
frmsize2	0.103 (0.046)**		
frmsize3	0.225 (0.068)***		
frmsize4	0.192 (0.064)***		
logearn	-0.010 (0.036)	0.104 (0.070)	-0.001 (0.063)
married		-0.111 (0.047)**	
pearn		-0.355 (0.319)	0.484 (0.306)
preex12	0.138 (0.049)***	0.087 (0.060)	
preex60cumst		0.023 (0.030)	0.020 (0.025)
preex60sq	0.050 (0.025)**	-0.028 (0.030)	
pretx1		-0.033 (0.125)	0.105 (0.122)
pretx2		0.252 (0.097)***	0.021 (0.101)
pretx35		0.103 (0.068)	0.239 (0.063)***
qual_h		0.133 (0.114)	
qual_h_a4055		0.578 (0.203)***	0.571 (0.169)***
qual_m		-0.026 (0.073)	
qual_m_a4055		0.571 (0.145)***	0.487 (0.104)***
y93q2		0.109 (0.110)	-0.011 (0.082)
y93q3		0.225 (0.103)**	0.058 (0.079)
y93q4		0.400 (0.098)***	0.104 (0.079)
y94q1		0.512 (0.096)***	0.124 (0.079)
y94q2	0.390 (0.055)***	0.507 (0.100)***	0.097 (0.084)

Propensity Score Estimates, SPST vs Waiting, Cohort 93/94 – continued			
	Stratum 1	Stratum 2	Stratum 3
y94q3	0.311 (0.056)***	0.567 (0.097)***	0.160 (0.081)**
y94q4	0.409 (0.052)***	0.554 (0.097)***	0.021 (0.084)
_cons	-2.308 (0.168)***	-2.229 (0.229)***	-1.913 (0.241)***
N	24751	14232	9913

Balancing Tests, Treatment SPST vs Waiting, Cohort 93/94				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	29	30	30	31
Stratum 2	42	44	44	45
Stratum 3	33	37	37	37
Quartic of Pscore				
Stratum 1	29	30	30	31
Stratum 2	37	40	43	45
Stratum 3	34	36	36	37

Propensity Score Estimates, Treatment RT vs Waiting, Cohort 93/94

	Stratum 1	Stratum 2	Stratum 3
BER2	0.312 (0.103)***	0.048 (0.124)	0.215 (0.145)
BER3	0.017 (0.107)	0.111 (0.127)	0.234 (0.152)
state11		-0.211 (0.106)**	
state1112	-0.174 (0.072)**		
state12		-0.210 (0.108)*	
industry3	-0.089 (0.108)		
industry4	-0.096 (0.122)		
industry5	-0.253 (0.129)*		
industry67	0.050 (0.092)		
a2529		1.028 (0.188)***	0.424 (0.114)***
a3034	-0.042 (0.069)	1.021 (0.181)***	0.605 (0.109)***
a3539		0.811 (0.181)***	0.559 (0.119)***
a3544	-0.556 (0.156)***		
a4044		0.612 (0.184)***	
a4549	-0.497 (0.135)***		
a5055	-0.822 (0.155)***		
urtb	0.019 (0.013)		
denst	-0.013 (0.030)	-0.083 (0.035)**	
f_age		0.016 (0.011)	

Propensity Score Estimates, RT vs Waiting, Cohort 93/94 – continued			
	Stratum 1	Stratum 2	Stratum 3
f_preex12		0.594 (0.197)***	
f_preex60cumst	-0.078 (0.055)	-0.285 (0.081)***	
f_qual_m	0.331 (0.133)**		
female	-0.216 (0.116)*	-1.137 (0.397)***	
for_age	0.015 (0.012)	-0.009 (0.003)**	
foreign	-0.951 (0.447)**		-0.227 (0.115)**
frmsize2	0.121 (0.068)*		
frmsize34	0.267 (0.078)***		
logearnsq			0.012 (0.010)
m_preex60cumst	-0.091 (0.047)*		
preex12	0.116 (0.072)		
preex24	0.181 (0.078)**		
preex60cumst		0.071 (0.046)	-0.043 (0.043)
pretx35			0.310 (0.103)***
qual_h		-0.462 (0.209)**	
qual_m	-0.251 (0.086)***		
qual_m_a3544	0.462 (0.166)***		
uentry	0.015 (0.014)	0.012 (0.018)	
uentry2	-0.001 (0.001)	-0.001 (0.001)	
y94q34			-0.260 (0.108)**
_cons	-2.722 (0.216)***	-3.049 (0.224)***	-2.976 (0.196)***
N	24421	13889	9350

Balancing Tests, Treatment RT vs Waiting, Cohort 93/94				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	25	26	26	27
Stratum 2	15	17	18	18
Stratum 3	8	10	10	10
Quartic of Pscore				
Stratum 1	23	23	25	27
Stratum 2	12	12	12	18
Stratum 3	7	9	10	10

Propensity Score Estimates, Treatment PF vs SPST, Cohort 93/94

	Stratum 1	Stratum 2	Stratum 3
BER2	0.634 (0.204)***		
BER3			-0.708 (0.207)***
state10	0.185 (0.248)	-0.552 (0.329)*	
state11	-0.101 (0.274)	-0.585 (0.260)**	
state12	0.232 (0.215)	-0.033 (0.207)	
state6	0.062 (0.283)	0.163 (0.267)	0.277 (0.207)
state7	0.743 (0.216)***	0.343 (0.212)	0.627 (0.162)***
state9	0.665 (0.244)***	0.439 (0.252)*	
a2529	-0.356 (0.162)**		
a3034		-0.535 (0.245)**	
a3555		0.056 (0.165)	
a4044			0.188 (0.174)
a4549			0.327 (0.203)
a5055			0.718 (0.194)***
urtb100	-5.813 (2.990)*		
denst		-0.222 (0.068)***	
f_BER2	-0.303 (0.330)		
f_BER3			1.665 (0.340)***
f_a3034		-0.272 (0.411)	
f_logearn		-2.979 (1.379)**	
f_logearnsq		0.547 (0.195)***	
f_preex12	-0.295 (0.312)		
f_preex24	0.196 (0.242)		
f_preex6	-0.063 (0.310)		
f_qual_u	-1.233 (0.455)***		
f_tnull			-0.006 (0.018)
female	0.671 (0.331)**	2.948 (2.640)	-7.643 (8.579)
foreign		0.602 (0.201)***	
logearn		1.923 (1.307)	
logearnsq		-0.320 (0.170)*	
m_tnull			-0.023 (0.012)*
preex60sq	-0.175 (0.084)**		
qual_u	0.523 (0.233)**		
y93q2	0.196 (0.271)	-0.561 (0.317)*	
y93q3	0.208 (0.263)		
y93q34		-0.458 (0.241)*	
y93q4	-0.118 (0.271)		
y94q1	0.206 (0.269)	-0.797 (0.267)***	
y94q2	-0.297 (0.255)	-0.989 (0.292)***	
y94q3	-0.205 (0.249)	-0.847 (0.269)***	
y94q4	-0.681 (0.268)**	-1.246 (0.293)***	

Propensity Score Estimates, PF vs SPST, Cohort 93/94 – continued			
	Stratum 1	Stratum 2	Stratum 3
_cons	-0.853 (0.399)**	-2.614 (2.603)	7.908 (4.722)*
N	630	583	755

Balancing Tests, Treatment PF vs SPST, Cohort 93/94				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	24	24	24	24
Stratum 2	22	22	22	22
Stratum 3	10	10	10	10
Quartic of Pscore				
Stratum 1	23	24	24	24
Stratum 2	21	22	22	22
Stratum 3	7	9	10	10

Propensity Score Estimates, Treatment PF vs RT, Cohort 93/94

	Stratum 1	Stratum 2	Stratum 3
BER3	0.476 (0.197)**	-0.540 (0.308)*	
state10		0.224 (0.440)	
state11	0.323 (0.330)	0.061 (0.361)	
state12	0.748 (0.246)***	0.698 (0.310)**	
state6	0.480 (0.340)	0.623 (0.360)*	
state7	0.615 (0.242)**	0.810 (0.275)***	
state9		0.844 (0.347)**	
state910	0.602 (0.247)**		
industry7		-0.724 (0.246)***	
a3034		-0.228 (0.257)	
a3539	0.281 (0.228)		
a3544		0.608 (0.239)**	
a4044	0.654 (0.224)***		
a4549	0.710 (0.408)*	1.763 (0.450)***	
a5055	0.412 (0.453)	2.046 (0.631)***	
f_BER3		1.427 (0.481)***	
f_BER34			1.923 (0.524)***
f_state10			-0.236 (1.087)
f_state11			-0.768 (1.137)
f_state12			0.781 (0.754)
f_state6			0.054 (0.959)

Propensity Score Estimates, PF vs RT, Cohort 93/94 – continued			
	Stratum 1	Stratum 2	Stratum 3
f_state7			0.448 (0.744)
f_state9			0.206 (0.898)
f_a3539			-0.410 (0.592)
f_a4044		0.976 (0.536)*	
f_a4055			1.568 (0.546)***
f_a4549	0.818 (0.652)		
f_married			-0.503 (0.466)
f_qual_m	1.091 (0.487)**		
female	-1.111 (0.481)**	-0.804 (0.330)**	-0.353 (0.926)
foreign	0.502 (0.307)	0.602 (0.276)**	
m_BER25			1.370 (0.408)***
m_state10			0.971 (0.659)
m_state11			0.165 (0.421)
m_state12			-1.058 (0.642)*
m_state6			0.460 (0.501)
m_state7			0.984 (0.407)**
m_state9			-0.066 (0.510)
m_a3539			0.955 (0.351)***
m_a4055			2.325 (0.485)***
m_married			-1.142 (0.384)***
tnull		-0.018 (0.015)	
_cons	-1.212 (0.191)***	6.714 (5.956)	-1.844 (0.511)***
N	300	240	192

Balancing Tests, Treatment PF vs RT, Cohort 93/94				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	13	13	13	14
Stratum 2	14	16	17	17
Stratum 3	20	20	21	21
Quartic of Pscore				
Stratum 1	13	14	14	14
Stratum 2	15	17	17	17
Stratum 3	17	20	20	21

Propensity Score Estimates, Treatment RT vs SPST, Cohort 93/94

	Stratum 1	Stratum 2	Stratum 3
BER2	0.808 (0.122)***		
state10	-0.512 (0.199)**	-0.384 (0.234)	
state11	-0.312 (0.186)*	-0.338 (0.194)*	
state12	-0.404 (0.151)***	-0.523 (0.185)***	
state79			0.278 (0.133)**
a2529		1.612 (0.505)***	
a3034		1.511 (0.507)***	
a3539	-0.241 (0.154)		
a3544			-0.323 (0.131)**
a3549		1.046 (0.502)**	
a4044	-0.286 (0.169)*		
a4549	-0.725 (0.233)***		
a4555			-0.799 (0.196)***
a5055	-0.896 (0.283)***		
f_BER2		0.743 (0.242)***	0.473 (0.215)**
f_married	-0.137 (0.179)		
f_preex60cumst		-0.180 (0.117)	
f_qual_m		0.382 (0.260)	
f_qual_u	-0.490 (0.297)*		
female	0.502 (0.163)***	-0.893 (0.303)***	-0.097 (0.164)
logearn			0.199 (0.127)
m_married	0.360 (0.148)**		
m_qual_h		-1.020 (0.359)***	
m_qual_m		-0.303 (0.196)	
married			0.354 (0.174)**
marriedBER2			-0.302 (0.217)
pearn			-0.658 (0.797)
qual_h			-0.458 (0.292)
qual_u	0.521 (0.183)***		
y93q2	-0.187 (0.250)	-0.427 (0.286)	
y93q3	0.237 (0.218)	-0.480 (0.265)*	
y93q4	-0.111 (0.224)	-0.630 (0.250)**	
y94q1	0.084 (0.220)	-0.567 (0.245)**	
y94q2	-0.215 (0.206)	-0.602 (0.270)**	
y94q3	-0.289 (0.214)	-0.933 (0.268)***	
y94q34			-0.340 (0.155)**
y94q4	-0.481 (0.209)**	-1.123 (0.270)***	
_cons	-0.900 (0.199)***	-0.949 (0.545)*	-1.088 (0.628)*
N	726	619	775

Balancing Tests, Treatment RT vs SPST, Cohort 93/94				
	P-values>.1	P-values>.05	P-values>.01	Regressors
Cubic of Pscore				
Stratum 1	20	20	20	20
Stratum 2	17	18	18	19
Stratum 3	10	11	11	11
Quartic of Pscore				
Stratum 1	17	18	20	20
Stratum 2	17	18	19	19
Stratum 3	11	11	11	11

A.1.3 Common Support

Figure A.1: Densities of Propensity Scores for Cohort 86/87

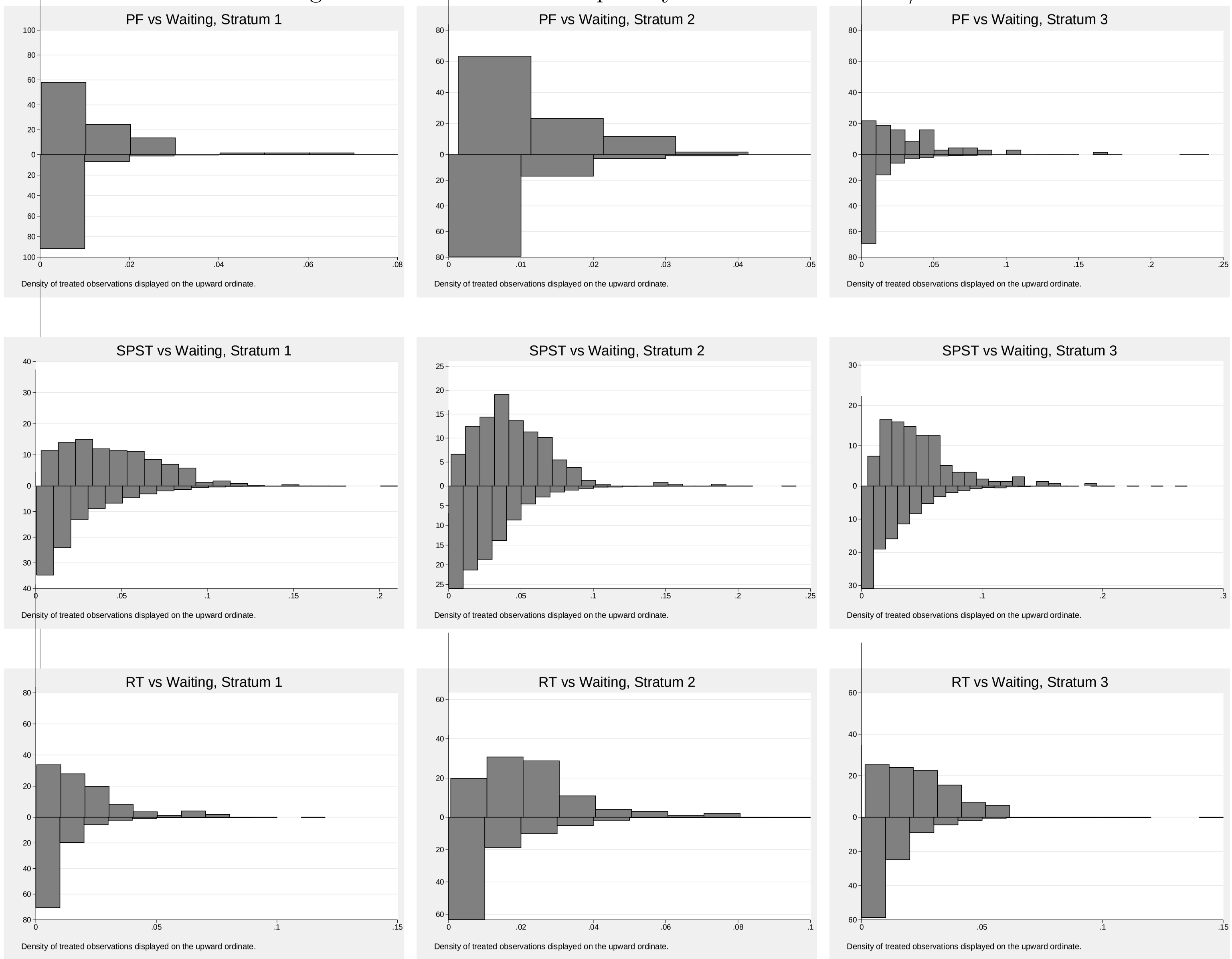


Figure A.2: Densities of Propensity Scores for Cohort 86/87

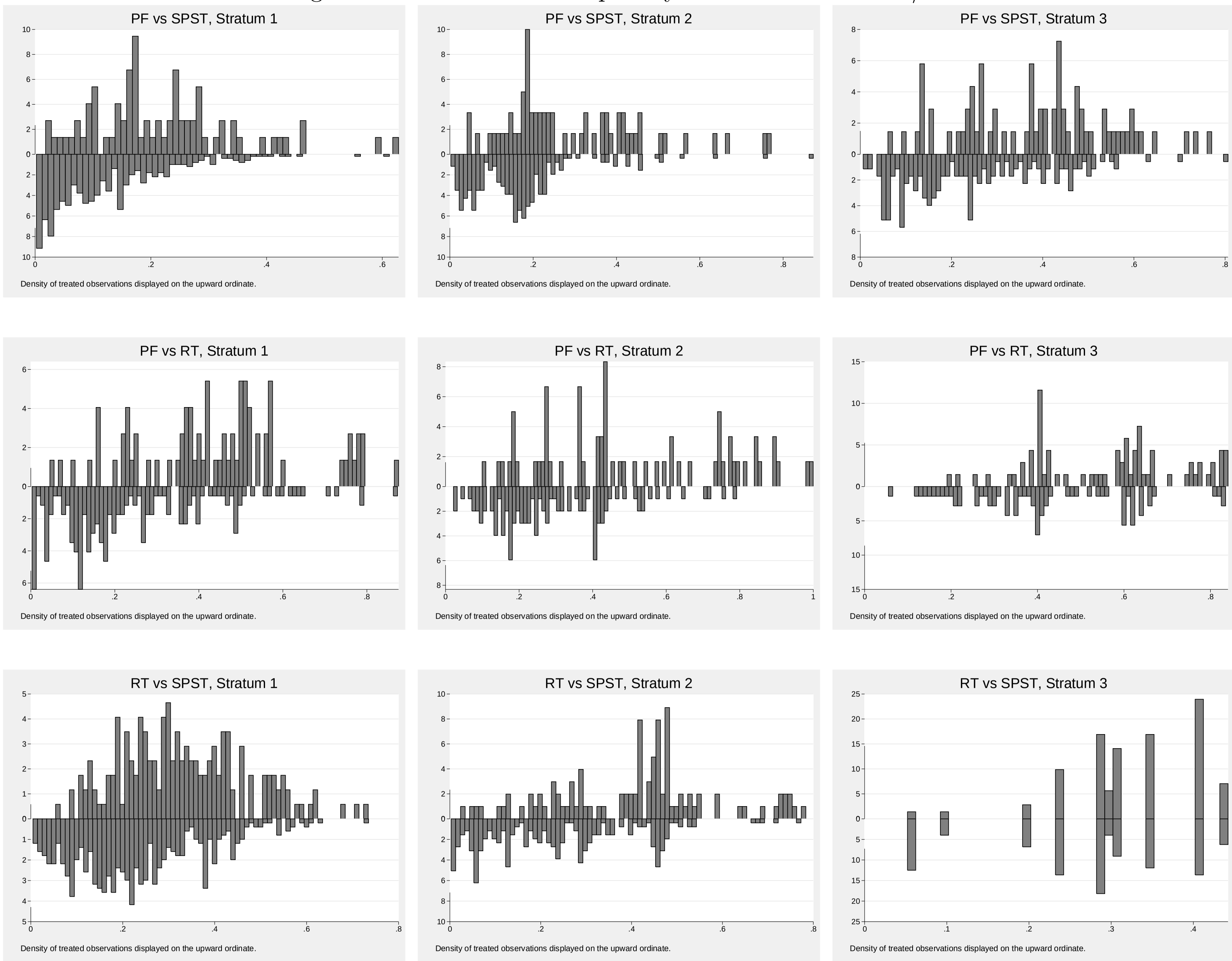


Figure A.3: Densities of Propensity Scores for Cohort 93/94

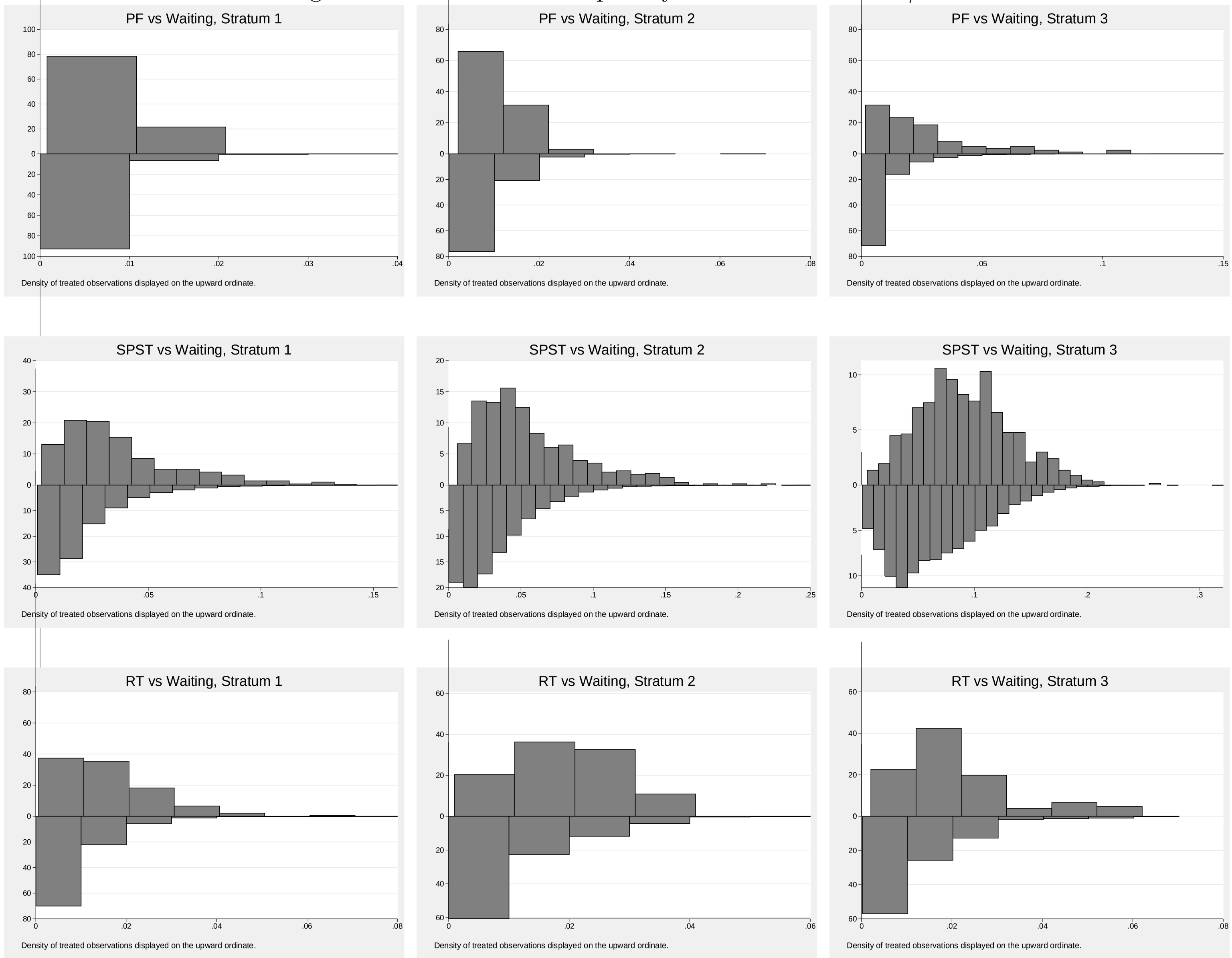
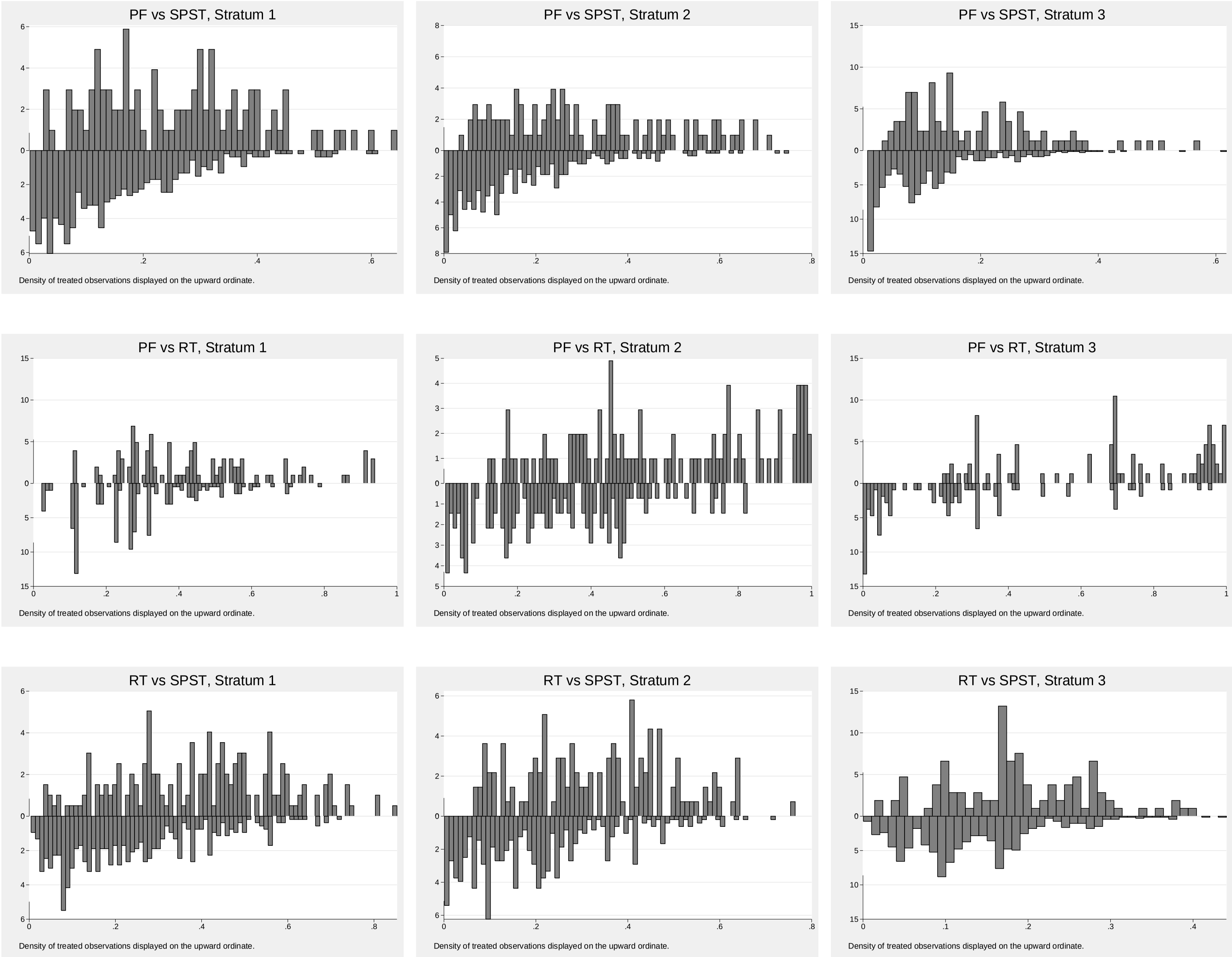


Figure A.4: Densities of Propensity Scores for Cohort 93/94



A.2 Background Information about the Data

Types of Further Training under the Labor Promotion Act

In this study we are interested in active labor market programs for unemployed who have previously been employed and who have not already found a new job. Here however, we want to give a short overview over the full set of training schemes administered under the Labor Promotion Act (Arbeitsförderungsgesetz, AFG).

Further Vocational Training (Berufliche Fortbildung)
A bunch of different training courses is subsumed under this heading. It includes theoretical as well as practical training schemes within the occupation of the participant.

Retraining (Umschulung)
This training scheme provides a complete, new vocational degree according to the German apprenticeship system.

Short–term Training according to §41a AFG
These programs last only about four weeks and were offered from 1979 until 1992. They are mainly intended to evaluate the participant's problems in finding regular employment. Starting 1993 such programs are no longer recorded as independent programs but as part of the regular counseling for unemployed. Hence we can not identify them in the inflow sample 1993/94. In order to make the samples comparable we treat the programs according to §41a in the 1986/87 inflow sample also as open unemployment. Thus if an unemployed first takes part in a §41a program and later in the same unemployment spell in Retraining we would consider the retraining the first program and evaluate it.

German Language Course (Deutschsprachlehrgang)
The German Language Courses are intended for newly arrived immigrants. So the participants typically have not been employed in Germany before the German Course and hence are not part of the focus group of this study, the previously employed unemployed.

Career Advancement (Aufstiegsförderung)
These programs are typically targeted at the employed and have been more important when the Labor Promotion Act was introduced in 1969. By providing

additional human capital the participant's risk of becoming unemployed should be lowered. Prime examples are courses in which the participants with a vocational training degree obtain additional certificates which allow them to independently run craftsman's establishments and to train trainees in the dual system of vocational training.

Integration Subsidy (Einarbeitungszuschuss)
Wage subsidies are paid for the employment of formerly long–term unemployed and are intended to decrease the competitive disadvantage of these recruits for the period of familiarization with the skill requirement of the job. Even if the target group of wage subsidies are also unemployed we do not evaluate them because they require a job for which the wage subsidy is paid. This means provision of wage subsidies is already conditional on employment which is the success criteria for the other programs.

Construction of the Monthly Panel

The IABS employment and LED benefit payment data are daily register data whereas the FuU training data gives monthly information about program participation. This study uses the merged data as described in Bender et al. (2005). From the merged data we construct a monthly panel. If the original daily data contain more than one spell overlapping a specific month we take the information from the spell with the largest overlap as the spell defining the monthly information.

The defining condition to be part of our inflow sample into unemployment is a transition from an employment month to a nonemployment month, in which the last employment month was between December 1985 (1992) and November 1986 (1993) and thus the first unemployment month was between January 1986 (1993) and December 1987 (1994). In order to divide nonemployment (to be precise: not employed subject to social security contributions) into unemployment and other states (like labor market leavers, transition into self employment, employment as civil servant) we additionally require a month with benefit payments from the employment office within the first twelve month of nonemployment or indication of participation in any labor market program in one of our data to be part of the inflow sample in unemployment.

Later on we aggregate the information further from monthly to quarterly information. Whereas the monthly employment information is binary the quarterly employment information can take the values 0, 1/3, 2/3 or 1.

We identify program participation if a person starts a program while being in the defining unemployment spell. The participant must not be employed in in the first month of the program. Otherwise we would consider such a program as a program which starts together with a job which we do not evaluate. In this case we would treat such a person as being employed. The exact identification of the program types will be explained in the following.

Identifying Program Participation from the Data

We identify participation in a further training program from a combination of FuU training data information, the benefit payment information and the employment status information. In principle, every participant in a further training program should be recorded in the FuU training data and we would not need the benefit payment data for identification of participation. There are two reasons to use the benefit payment data as well. First, we find the training data to be incomplete, many recipients of training related benefits are not contained in the training data.[1] Only using the benefit payment data identifies these participants. Second, quite often the type of training in the training data is given very unspecific as "Other adjustment of working skills". The benefit payment data can give more information about these programs. Finally, we need the employment status to identify participation because we only evaluate programs which start while being unemployed. In particular we do not consider integration subsidies which are associated with a regular job. We exclude programs starting together with a job because our outcome variable is employment and program starts that are conditional on having found a job are (partly) endogenous.

In the remaining part of this section we describe how we aggregated the benefit payment information and the training data information. The next section contains the exact coding plan. We disclose in detail which combination of information from benefit payment and training data we identify as PF, SPST or RT.[2]

[1] Remember the purpose of the training data was only internal documentation. This might explain its incompleteness.

[2] More details about the benefit payment data and training data can be found in Speckesser (2004), Fitzenberger and Speckesser (2007) and Bender et al. (2005).

Benefit Payment Information in the LED–Data

The merged data we use contain three variables with benefit payment information from the original LED data, ("parallel original benefit information 1–3" [*Leistungsart im Original 1–3*] L1LA1, L2LA1, L3LA1). The main variable is L1LA1. If there are two parallel payment informations in the original data L1LA2 also contains information and only if there is a third parallel payment spell L3LA1 is also filled. In general we use L1LA1. Only if L1LA1 is not informative about program participation and L2LA1 is we use L2LA1 and only if L1LA1 and L2LA1 are not informative but L3LA1 we use L3LA1. The benefit payment information is given in time varying three–digit codes (for the coding plan see Bender et al. 2005). We extracted the program related information from the benefit payment information as given in table B.20. The main distinction regarding program participation is the distinction between no benefits at all or unemployment benefits/assistance on the one hand and program related maintenance benefits on the other hand. There are five types of program related benefits. Most important for us are the more general maintenance benefits while in further training and the more specific maintenance benefits while in retraining.

Table A.11: Aggregated types of benefit payment

German Abbreviation	Description
ALG	Unemployment benefits
ALHi	Unemployment assistance
UHG §41a	Income maintenance while in specific short term training program
UHG Fortbildung	Income maintenance while in further training
UHG Umschulung	Income maintenance while in retraining
UHG Darlehen	Income maintenance as a loan
UHG Deutsch	Income maintenance while in a German course

The original benefit payment information is given in three variables L1LA1, L2LA1 and L3LA1 with time varying three–digit codes.

Training Types in the FuU–Data

In this evaluation study one of the most important advantages compared to survey data is the information about the precise type of training. It allows us to identify homogeneous treatments for the evaluation. In the merging process, up to two parallel FuU–spells were merged to one spell of the IABS data because in many cases the FuU–data provided more than one parallel spell. These two parallel spells provide two variables indicating the type of course (*Maßnahmeart* [FMASART1, FMASART2]).

Correcting Type of Training for 1986 The annual frequency for the type of training *14* in 1986 looks very different than in the years before and after. Additionally the distributions of the planned durations and the types of examination completing the program *14* in 1986 are different than in the adjacent years. We think this is due to a lacking recoding of *14* to *12*, which was necessary for the years until 1985 because the coding of FMASART changed over the years. Hence we recode *14* to *12* in 1986 if the planned duration is less than 10 month.

Aggregating the training type information Since type of treatment (*Maßnahmeart*) is often coded as "other adjustment"
(FMASART1=12 [*Sonstige Anpassungen*]) in the FuU–data, we increase the precision of information about the type of treatment by relying on the second parallel information about the type of training: The second FuU–spell is used if the first FuU–spell is coded as "other adjustment" (*"Sonstige Anpassungen"*) and a second spell includes a code different from 12. Such combined information of FMASART1 and FMASART2 is referred to as FMASART* in the following.

Combining the Information

When using information from different sources, the sources may give differing information. If the training data indicated training participation and the benefit payment data did not or vice versa we relied on the source which indicated training for the following reasons. If somebody receives training related benefits it is more likely that the employment agency forgot to fill in the training data record than the agency wrongly induced payment of benefits. And if somebody is contained in the training data but does not receive maintenance benefits he either receives no benefits, which is possible while being in training, or receives unemployment benefits/assistance and the payment is just wrongly labeled.

If both training and benefit payment data indicate program participation but differ in the type of program we generally use the training data information. An example: the benefit payment indicates maintenance payments for further training and the training data indicates Retraining. We use Retraining from the training data.

The only exception is unspecific program information from the training data "other adjustment". If in such cases the benefit payment data give specific information like Retraining we use the information from benefit payment data. All possible combinations of training and benefit payment information which we use to identify participation in one of the three programs are given in the following section.

Coding Plan for the Treatment Information

This section gives the exact coding plans for identification of Practice Firm, SPST and Retraining. In general we identify program participation as start of a program in an unemployment spell before another employment begins. This means that we only identify a start of a program if the employment status in the first month of the program indicates no employment (BTYP≠1).

Practice Firm

Practice Firm is a consolidation of the program types Practice enterprise and Practice studio from the FuU training data. There is no specific benefit payment type related to Practice Firms, rather the participants shall receive the general maintenance payment for further training. Since the training data are more reliable than the benefit payment data regarding type of the program we identify Practice Firm whenever FMASART shows the codes 11 or 12 independently of the payment information.

Program code	Label	Label in German
10	Practice enterprise	Übungsfirma
11	Practice studio	Übungswerkstatt

In table B.21 we show how often which combination of benefit payment information and program type information identifies *Practice Firm* in the two inflow samples.

Table A.12: Identification of *Practice Firm* with program type and benefit payment type: Frequencies

Program	Type of payment No benefits	UB/UA	Income Maintenance for Short–term Training	Further Training	Retraining	Total
Practice enterprise	4	5	1	198	2	210
Practice studio	11	19	0	311	20	361
Total	15	24	1	509	22	571

Both inflow samples together. BTYP≠1 as an additional requirement.

Provision of Specific Professional Skills and Techniques

We identify SPST in the following cases.

(a) Identification from training data and benefit payment data
We identify SPST if the training data indicates the general program "Other adjustment" and the benefit payment information is no benefit payments, unemployment benefits, unemployment assistance or maintenance payments while in retraining.

Program code	Label	Label in German
12	Other adjustment of working skills	sonst. Anpassung der berufl. Kenntnisse

(b) Reliance on benefit payment data
We identify SPST if the program information from the training data is missing and the benefit payment information is maintenance payments while in further training.

Program code	Label	Label in German
-9	missing	fehlende Angabe

(c) Additional program from training data
We also identify SPST when another program of little quantitative importance but comparable content is recorded in the training data independent of the benefit payment information.

Program code	Label	Label in German
31	Further education of trainers and multidisciplinary qualification	Heran-/Fortbildung v. Ausbildungskräften/ berufsfeldübergreifende Qualifikation

(d) Additional combinatioin

Finally we identify SPST if the training data indicate the unspecific "other career advancement" and the benefit payment information indicates further training.

Program code	Label	Label in German
28	Other promotion	sonstiger Aufstieg (< 97)

In table B.22 we show how often which combination of benefit payment information and program type information identifies *SPST* in the two inflow samples.

Table A.13: Identification of *SPST* with program type and benefit payment type: Frequencies

Program	Type of payment No benefits	UB/UA	Income Maintenance for further training	Total
Missing	0	0	644	644
Other adjustment of working skills	57	89	2095	2241
Other promotion	0	0	150	150
Further education of trainers and multidisciplinary qualification	0	1	1	2
Total	57	90	2890	3037

Both inflow samples together. BTYP≠1 as an additional requirement.

Retraining

Retraining or longer "Qualification for the first labor market via the education system" is taking part in a new vocational training and obtaining a new vocational training degree according to the German dual education system. Additionally, but quantitatively of little importance we see the make up of a missed examination "Certification" as comparable to retraining because the result is the same. Furthermore and also only of marginal importance we see participation in the programs "Technican" or "Master of Business administration (not comparable to an american style

MBA)" while not receiving maintenance benefits as a loan as Retraining. Conventionally these two programs are considered as career advancement programs which we do not evaluate. Benefits as a loan would underline their character as career advancements.

(a) Identification from training data

We identify the following two programs as Retraining independent of the benefit payment information.

Program code	Label	Label in German
29	Certification	berufl. Abschlussprüfung
32	Retraining	Umschulung

(b) Reliance on benefit payment data

If the training data is uninformative and maintenance benefits for Retraining are paid we identify Retraining.

Program code	Label	Label in German
-9	missing	fehlende Angabe
12	Other adjustment of working skills	sonst. Anpassung der berufl. Kenntnisse

(c) Other programs from training data

Two other programs are identified from the training data. They typically also take two years full time and require an existing vocational training degree, hence are somewhat comparable to retraining in a narrower definition. Not identified if maintenance benefits are paid as a loan.

Program code	Label	Label in German
26	Technician	Techniker (<97)
27	Master of business administration	Betriebswirt (<97)

In table B.23 we show how often which combination of benefit payment information and program type information identifies *Retraining* in the two inflow samples.

Table A.14: Identification of *Retraining* with program type and benefit payment type: Frequencies

Program	Type of payment					Total
	No benefits	UB/UA	Income Maintenance			
			Further Training	Retraining	Loan	
missing	0	0	0	110	0	110
Other adjustment of working skills	0	0	0	65	0	65
Technician	2	1	5	2	0	10
Master of business administration	0	2	1	1	0	4
Certification	4	1	20	7	0	32
Retraining	11	13	231	355	2	612
Total	17	17	257	540	2	833

Both inflow samples together. BTYP≠1 as an additional requirement.

Appendix B

Additional Appendix to Chapter 2

B.1 Estimation Results for the Propensity Score

B.1.1 Sample Sizes and Variable Definitions

Sample Sizes by Stratum

	Women		
	Stratum 1	Stratum 2	Stratum 3
Waiting	5783 (4652)	3855 (2996)	2294 (1671)
PF	37	51	48
SPST	254	374	435
RT	61 (61)	76 (75)	53 (53)
	Men		
	Stratum 1	Stratum 2	Stratum 3
Waiting	5558 (4444)	2705 (2046)	1381 (997)
PF	40	15	14
SPST	200	141	144
RT	113 (107)	82 (79)	35 (33)

Remark: Numbers in Parentheses exclude the 51–55 year old. We use this further restricted sample to evaluate RT. We do not evaluate PF for males in stratum 2 and 3 due to the small sample size.

Sample Sizes by Quarter

Quarter of unempl.	Inflows	Outflows Job	PF	SPST	RT	controls	share trt later	alternative controls	alt share trt later
				Women					
1	6135	621	12	119	33	5783	0.207	5971	0.232
2	5350	806	25	135	28	5162	0.232	5162	0.232
3	4356	582	28	206	42	3855	0.181	4080	0.226
4	3498	443	23	168	34	3273	0.213	3273	0.213
5	2830	408	21	157	31	2294	0.070	2621	0.186
6	2213	241	10	115	11	1886	0.085	2077	0.169
7	1836	155	10	99	7	1645	0.098	1720	0.137
8	1565	127	7	64	4	1490	0.108	1490	0.108
9+	1363	1202	9	147	5		mean		mean
							0.175		0.206
				Men					
1	5911	1315	20	80	57	5558	0.087	5754	0.118
2	4439	1300	20	120	56	4243	0.114	4243	0.114
3	2943	729	10	84	52	2705	0.090	2797	0.120
4	2068	402	5	57	30	1976	0.123	1976	0.123
5	1574	272	4	52	23	1381	0.037	1495	0.110
6	1223	176	2	34	6	1109	0.046	1181	0.104
7	1005	102	3	29	3	933	0.055	970	0.091
8	868	68	5	29	3	831	0.061	831	0.061
9+	763	712	4	43	4		mean		mean
							0.088		0.113

Remark: The table shows quarter by quarter of elapsed unemployment duration the number of those who are still unemployed at the beginning of the quarter (inflows) and the number of those who during the quarter start a job (job) or a treatment (PF, SPST, RT). Controls are all those who are still unemployed at the beginning of the quarter but do not start a treatment during the stratum. The share of the controls who start a treatment during a later stratum is also given. An alternative definition of control persons (not pursued in the paper) would take as controls all those, who are still unemployed at the beginning of the quarter but do not start a treatment during the *quarter*. This would lead to a slightly higher share of controls who receive treatment later. The means are weighted means. The table considers the sample age 25–55 at the beginning of unemployment. The restricted version age 25–50 for RT is available upon request from the authors. The number for outflows in jobs in quarter 9+ include those, who never again start a job.

Variable Definitions

Table B.1: Variable Definitions

Label	Definition
Personal Attributes	
aXXYY	Age at start of unemployment $\geq$XX and $\leq$ YY
age	Age at start of unemployment
married	Married
qual_l	No vocational training degree or education information missing
qual_m	Vocational training degree
qual_h	University/College degree
Last Employment	
BER1	Apprentice
BER2	Blue Collar Worker
BER3	White Collar Worker
BER4	Worker at home with low hours or BER missing
BER5	Part–time working
pearn	Daily earnings $\geq$ 15 Euro per day in 1995 Euro
earnlow	Daily earnings $<$ 15 Euro per day in 1995 Euro
earncens	Earnings censored at social security taxation threshold
earn	Daily earnings
logearn	log(earn) if pearn=1 and earncens=0, otherwise zero
Last Employer	
industry1	Agriculture
industry2	Basic materials
industry3	Metal, vehicles, electronics
industry4	Light industry
industry5	Construction
industry6	Production oriented services, trade, banking
industry7	Consumer oriented services, organization and social services
frmsize1	Firm Size (employment) missing or $\leq$ 10
frmsize2	Firm Size (employment) $>$ 10 and $\leq$ 200
frmsize3	Firm Size (employment) $>$ 200 and $\leq$ 500

<continued on next page>

Table B.1: Variable Definitions <continued>

Label	Definition
frmsize4	Firm Size (employment) > 500
Employment and Program History	
preexM	Employed M (M=6, 12) month before unemployment starts
preex12cum	Number of months employed in the last 12 months before unemployment starts, standardized
pretx1	Participation in any ALMP program reported in our data in the year before unemployment starts
Regional Information	
state1	Mecklenburg-Vorpommern
state2	Brandenburg
state3	Sachsen-Anhalt
state4	Sachsen
state5	Thüringen
popdens	population density (standardized)
Calendar Time of Entry into Unemployment	
uentry	First unemployment month (months counted from January 1993
Interaction of Variables / Functional Form	
_sq	squared
_	interaction

All variables are defined at the time of entry into unemployment and constant during the unemployment spell.

B.1.2 Summary Statistics

The following six tables document the mean values of the variables in the three strata for women and men. The means are shown for the dynamic control group and the participants in PF, SPST and RT, respectively. Since we restrict the age for the evaluation of RT to lie between 25 and 50, we show the means for the dynamic control group also for this more restricted group and for RT only for this age group.

Table B.2: **Women Stratum 1**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	40.088	36.886	40.757	39.819	33.689
married	.59	.576	.514	.61	.738
qual l	.108	.098	.081	.039	.066
qual m	.843	.851	.892	.89	.869
qual h	.049	.05	.027	.071	.066
BER1	.003	.004	0	0	0
BER2	.428	.425	.432	.252	.443
BER3	.391	.399	.405	.535	.393
BER4	.002	.002	0	0	0
BER5	.176	.171	.162	.213	.164
pearn	.966	.967	.973	.996	1
earncens	.004	.004	0	.004	0
logearn	3.447	3.445	3.521	3.647	3.657
industry1	.066	.064	.027	.047	.016
industry2	.047	.044	.027	.043	.115
industry3	.064	.064	.189	.075	.066
industry4	.074	.072	.081	.071	.033
industry5	.036	.038	.081	.035	.049
industry6	.264	.269	.162	.303	.18
industry7	.45	.45	.432	.425	.541
frmsize1	.229	.235	.081	.193	.18
frmsize2	.447	.447	.676	.433	.344
frmsize3	.152	.148	.189	.177	.213
frmsize4	.173	.169	.054	.197	.262
preex6	.842	.835	.784	.846	.934
preex12	.764	.747	.757	.827	.836
preex12cum	10.153	10.052	9.811	10.354	11.033
pretx1	.077	.083	.054	.031	.066
state1	.139	.144	.081	.154	.131
state2	.167	.169	0	.118	.246
state3	.209	.207	.162	.205	.115
state4	.304	.299	.216	.362	.377

Continued on next page...

... table B.2 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
state5	.18	.179	.541	.161	.131
popdens	468.023	463.781	422.692	544.084	379.189
uentry	10.721	10.847	8.892	13	10.033
N	5783	4652	37	254	61

Table B.3: **Women Stratum 2**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	40.777	37.183	41.176	38.623	33.12
married	.587	.568	.667	.61	.627
qual l	.133	.122	.039	.078	.08
qual m	.828	.838	.863	.834	.893
qual h	.039	.04	.098	.088	.027
BER1	.003	.003	0	.005	0
BER2	.432	.433	.353	.294	.427
BER3	.373	.378	.49	.559	.413
BER4	.002	.002	0	0	0
BER5	.191	.185	.157	.142	.16
pearn	.963	.963	1	.979	1
earncens	.003	.003	.02	.003	.013
logearn	3.427	3.419	3.562	3.593	3.609
industry1	.058	.058	0	.037	.053
industry2	.051	.048	.078	.029	.053
industry3	.061	.058	.098	.091	.08
industry4	.073	.073	.02	.056	.027
industry5	.038	.041	0	.027	.027
industry6	.261	.263	.294	.342	.32
industry7	.457	.459	.51	.417	.44
frmsize1	.222	.228	.176	.203	.173
frmsize2	.439	.439	.353	.428	.44
frmsize3	.156	.154	.235	.144	.16

Continued on next page...

... table B.3 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
frmsize4	.182	.179	.235	.225	.227
preex6	.859	.85	.922	.832	.92
preex12	.793	.77	.784	.786	.867
preex12cum	10.374	10.242	10.843	10.257	11.027
pretx1	.072	.079	.059	.053	.08
state1	.135	.138	.157	.139	.227
state2	.172	.174	.02	.155	.147
state3	.216	.215	.333	.233	.147
state4	.292	.288	.275	.329	.36
state5	.185	.186	.216	.144	.12
popdens	466.479	461.806	520.108	561.464	392.724
uentry	10.321	10.433	8.157	13.024	11.733
N	3855	2996	51	374	75

Table B.4: **Women Stratum 3**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	41.668	37.309	43.604	40.609	33.453
married	.589	.57	.542	.591	.547
qual l	.159	.147	.083	.099	.075
qual m	.805	.814	.896	.857	.906
qual h	.036	.038	.021	.044	.019
BER1	.003	.003	.021	0	0
BER2	.449	.454	.292	.368	.528
BER3	.341	.342	.583	.446	.434
BER4	.001	.002	0	.002	0
BER5	.206	.199	.104	.184	.038
pearn	.959	.959	.938	.982	1
earncens	.004	.003	0	.002	0
logearn	3.395	3.388	3.452	3.562	3.674
industry1	.06	.06	.042	.048	.038

Continued on next page...

... table B.4 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
industry2	.051	.046	.125	.064	.075
industry3	.063	.059	.104	.057	.094
industry4	.075	.073	.042	.094	.075
industry5	.041	.045	0	.037	.038
industry6	.246	.247	.313	.308	.189
industry7	.464	.47	.375	.391	.491
frmsize1	.221	.23	.188	.152	.075
frmsize2	.436	.437	.438	.453	.396
frmsize3	.158	.151	.063	.182	.17
frmsize4	.185	.181	.313	.214	.358
preex6	.866	.855	.917	.906	.887
preex12	.804	.774	.813	.853	.849
preex12cum	10.448	10.284	10.979	10.906	10.755
pretx1	.06	.068	.083	.078	.057
state1	.121	.121	.229	.103	.189
state2	.167	.168	.021	.2	.094
state3	.228	.228	.229	.172	.208
state4	.289	.284	.333	.368	.226
state5	.195	.199	.188	.156	.283
popdens	453.54	451.059	642.06	559.151	534.315
uentry	10.242	10.373	9.458	9.874	6.755
N	2294	1671	48	435	53

Table B.5: **Men Stratum 1**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	40.124	36.84	36.975	40.035	33.449
married	.466	.435	.475	.56	.421
qual l	.094	.089	.125	.06	.075
qual m	.842	.851	.825	.75	.897
qual h	.063	.059	.05	.19	.028

Continued on next page...

... table B.5 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
BER1	.001	.002	.025	0	0
BER2	.794	.804	.85	.6	.86
BER3	.167	.161	.075	.34	.112
BER4	0	0	0	.005	0
BER5	.037	.033	.05	.055	.028
pearn	.991	.991	1	.995	.991
earncens	.015	.014	0	.03	0
logearn	3.647	3.649	3.72	3.711	3.727
industry1	.079	.074	.05	.05	.084
industry2	.072	.075	.075	.115	.103
industry3	.116	.113	.1	.165	.196
industry4	.056	.056	.05	.05	.065
industry5	.207	.227	.175	.115	.121
industry6	.221	.234	.15	.29	.262
industry7	.248	.222	.4	.215	.168
frmsize1	.255	.27	.175	.19	.131
frmsize2	.501	.503	.425	.49	.533
frmsize3	.12	.109	.1	.115	.121
frmsize4	.124	.118	.3	.205	.215
preex6	.845	.835	.875	.845	.869
preex12	.776	.761	.75	.87	.776
preex12cum	10.223	10.097	10.175	10.56	10.533
pretx1	.064	.069	.025	.03	.075
state1	.14	.141	.2	.2	.168
state2	.147	.152	.05	.13	.14
state3	.209	.213	.4	.165	.243
state4	.323	.319	.15	.35	.346
state5	.18	.175	.2	.155	.103
popdens	460.115	463.21	340.144	604.944	499.402
uentry	10.6	10.657	9.425	12.01	10.533
N	5558	4444	40	200	107

Table B.6: **Men Stratum 2**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	41.316	37.473	38.933	40.099	34.443
married	.44	.396	.333	.44	.405
qual l	.12	.118	.133	.078	.114
qual m	.805	.812	.867	.801	.797
qual h	.075	.07	0	.121	.089
BER1	.002	.002	0	0	0
BER2	.761	.771	1	.645	.823
BER3	.197	.19	0	.319	.165
BER4	0	0	0	0	0
BER5	.04	.037	0	.035	.013
pearn	.988	.988	1	.993	1
earncens	.019	.016	0	.021	.013
logearn	3.61	3.615	3.67	3.666	3.715
industry1	.06	.053	.2	.035	.051
industry2	.078	.081	0	.057	.089
industry3	.115	.107	.067	.128	.165
industry4	.053	.055	.133	.043	.076
industry5	.174	.195	.133	.149	.139
industry6	.23	.242	.133	.312	.241
industry7	.29	.266	.333	.277	.241
frmsize1	.229	.247	.133	.199	.19
frmsize2	.487	.489	.6	.447	.57
frmsize3	.131	.118	.2	.163	.139
frmsize4	.153	.146	.067	.191	.101
preex6	.839	.82	.867	.844	.886
preex12	.783	.761	.733	.787	.861
preex12cum	10.216	10.005	10.067	10.277	10.772
pretx1	.062	.07	.067	.043	.038
state1	.134	.135	.333	.113	.19
state2	.152	.157	0	.106	.114
state3	.232	.236	.333	.284	.177
state4	.308	.304	.133	.355	.304

Continued on next page...

... table B.6 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
state5	.174	.167	.2	.142	.215
popdens	487.763	491.47	361.654	646.669	410.726
uentry	10.461	10.605	13.267	13.376	9.405
N	2705	2046	15	141	79

Table B.7: **Men Stratum 3**

Variable	control	control 25–50	PF	SPST	RT 25–50
age	42.142	37.872	41.929	42.201	34.364
married	.433	.391	.5	.41	.333
qual l	.125	.123	.071	.153	.242
qual m	.791	.794	.929	.674	.727
qual h	.084	.082	0	.174	.03
BER1	.001	.002	0	.007	0
BER2	.736	.741	.714	.611	.788
BER3	.218	.217	.286	.333	.182
BER4	0	0	0	0	0
BER5	.045	.04	0	.049	.03
pearn	.986	.987	1	.986	.97
earncens	.025	.022	0	.056	0
logearn	3.572	3.579	3.651	3.503	3.669
industry1	.056	.045	.071	.021	.03
industry2	.078	.077	.071	.069	.091
industry3	.122	.113	.143	.125	.091
industry4	.053	.06	.071	.035	.091
industry5	.152	.165	.071	.167	.182
industry6	.227	.237	.357	.236	.273
industry7	.311	.302	.214	.347	.242
frmsize1	.236	.259	.429	.16	.182
frmsize2	.461	.458	.5	.528	.455
frmsize3	.133	.116	0	.125	.182

Continued on next page...

... table B.7 continued

Variable	control	control 25–50	PF	SPST	RT 25–50
frmsize4	.169	.166	.071	.188	.182
preex6	.857	.839	.857	.861	.788
preex12	.789	.767	.786	.819	.788
preex12cum	10.319	10.129	9.857	10.493	9.667
pretx1	.052	.055	.071	.063	.121
state1	.127	.128	.357	.146	.394
state2	.165	.176	0	.153	.061
state3	.217	.221	.214	.229	.333
state4	.322	.312	.071	.313	.091
state5	.169	.163	.357	.16	.121
popdens	508.681	505.011	311.301	555.184	460.436
uentry	10.345	10.395	10.5	10.833	8.758
N	1381	997	14	144	33

B.1.3 Results of Propensity Score Estimations and Balancing Tests

Remark: The propensity score tables show the estimated coefficients of the probit regressions of the conditional probability to participate in the program mentioned in the header against the alternative of not taking part in any program in the stratum. The estimations are carried out separately for each time window of elapsed unemployment duration (Stratum 1, 2, and 3). Standard errors are in parentheses. *, **, *** means significant at the 10%, 5%, 1% level, respectively, in a two–sided test. Each probit table is followed by a table indicating how many regressors pass the Smith/Todd (2005) balancing test at different significance levels using a cubic and a quartic of the propensity score, respectively. Graphs with the densities of the propensity scores are in the next subsection.

Table B.8: **Propensity Score Estimates Women Practice Firm**

COEFFICIENT	Stratum 1		Stratum 2		Stratum 3	
age	0.0669	(0.073)	0.0972	(0.065)	0.0615	(0.075)
age_sq	-0.000798	(0.00089)	-0.00118	(0.00080)	-0.000643	(0.00089)
married	-0.221*	(0.13)			-0.169	(0.13)
qual_l	0.254	(0.43)				
qual_m	0.255	(0.39)				
BER2	-0.0267	(0.13)			-0.0419	(0.20)
entglow	0.292	(0.55)				
logearn_sq	0.0298	(0.029)	0.0213	(0.018)		
frmsize2	0.521***	(0.17)				
frmsize3	0.457**	(0.21)				
state3	0.399*	(0.24)			0.796**	(0.36)
state4	0.337	(0.22)	0.707**	(0.33)	0.840**	(0.36)
state5	0.901***	(0.21)			0.779**	(0.37)
BER3_married			0.193	(0.12)		
preex12			-0.0887	(0.14)	-0.0445	(0.16)
state1			0.801**	(0.34)	1.084***	(0.37)
state35			0.870***	(0.32)		
popdens			0.0436	(0.11)		
popdens_sq			-0.0213	(0.064)		
BER3					0.333*	(0.19)
pearn					-1.260*	(0.74)
logearn					0.283	(0.19)
uentry_sq					-0.000311	(0.00040)
Constant	-5.174***	(1.53)	-5.103***	(1.36)	-3.966**	(1.56)
Observations	5820		3906		2342	
PseudoR^2	0.0977		0.0397		0.0685	

Table B.10: **Propensity Score Estimates Women SPST**

COEFFICIENT	Stratum 1		Stratum 2		Stratum 3	
a3034	0.0240	(0.11)	0.123	(0.098)	0.247**	(0.11)
a3539	0.482*	(0.26)	0.107	(0.10)	-0.0693	(0.14)
a4044	0.523**	(0.26)	-0.0184	(0.10)	0.0334	(0.14)
a4549	0.124	(0.30)	-0.223*	(0.12)	0.0154	(0.14)
a5055	0.154	(0.30)	-0.310***	(0.10)	-0.288**	(0.13)

married	-0.00535	(0.063)	0.0481	(0.059)	-0.225**	(0.11)
qual_l	-0.178*	(0.10)	-0.0439	(0.084)	-0.199**	(0.085)
BER2	-0.313***	(0.089)	-0.0569	(0.088)	-0.0701	(0.086)
BER3	-0.0116	(0.087)	0.250***	(0.089)	0.0820	(0.094)
logearn_sq	0.0529***	(0.012)				
earncens	0.510	(0.51)	0.842	(0.68)	0.822	(0.70)
state1	-0.0397	(0.094)			-0.184*	(0.10)
state2	-0.248**	(0.098)			0.0280	(0.091)
state3	-0.0977	(0.085)			-0.282***	(0.088)
state5	-0.152*	(0.091)			-0.214**	(0.092)
uentry	-0.0546***	(0.015)	0.113***	(0.017)		
uentry_sq	0.00324***	(0.00063)	-0.00339***	(0.00069)		
preex12	0.652***	(0.21)				
preex12_a3544	-0.607**	(0.29)				
preex12_a4555	-0.748**	(0.35)				
preex12cum_sq	-0.0601	(0.071)	0.0410***	(0.015)	0.0333	(0.041)
preex12cum_sq_a3544	0.0320	(0.095)				
preex12cum_sq_a4555	0.318***	(0.10)				
preex12cum	-0.244*	(0.15)			0.168**	(0.076)
preex12cum_a3544	0.132	(0.20)				
preex12cum_a4555	0.636***	(0.24)				
pretx1	-0.453***	(0.16)	-0.286**	(0.13)	0.333***	(0.12)
qual_h			0.227*	(0.13)	-0.276	(0.17)
pearn			-1.080***	(0.40)	-0.765*	(0.43)
logearn			0.350***	(0.10)	0.307***	(0.11)
industry3			0.364**	(0.15)	-0.153	(0.15)
industry4			0.0907	(0.16)	0.117	(0.14)
industry5			-0.149	(0.20)	-0.102	(0.18)
industry67			0.112	(0.11)		
industry6					0.0405	(0.11)
industry7					-0.175*	(0.10)
frmsize2					0.247***	(0.087)
frmsize3					0.259**	(0.11)
frmsize4					0.179*	(0.11)
popdens					0.0718	(0.063)
popdens_sq					0.00989	(0.037)
married_a3544					0.318**	(0.16)
married_a4555					0.271*	(0.15)

Constant	-2.601***	(0.25)	-2.445***	(0.24)	-1.346***	(0.25)
Observations	6037		4229		2729	
PseudoR^2	0.0742		0.0807		0.0550	

Table B.12: **Propensity Score Estimates Women RT**

COEFFICIENT	Stratum 1		Stratum 2		Stratum 3	
a2529	0.597***	(0.18)			1.106***	(0.37)
a3034	0.404**	(0.18)			1.206***	(0.36)
a3539	0.0882	(0.21)			0.942**	(0.37)
a4044	0.0421	(0.21)			0.818**	(0.38)
married	0.377***	(0.12)	0.172	(0.11)	-0.0573	(0.14)
qual_l	-0.159	(0.17)	-0.116	(0.16)	-0.731**	(0.29)
qual_h	-0.0540	(0.25)	-0.307	(0.32)		
BER2	0.0616	(0.15)			0.592**	(0.29)
BER3	-0.100	(0.17)			0.542*	(0.30)
earn	0.0106**	(0.0042)	0.0110***	(0.0038)		
preex12cum	0.142**	(0.064)	0.263*	(0.14)		
state1	0.211	(0.18)				
state24	0.375***	(0.12)				
uentry	-0.0989***	(0.025)	0.133***	(0.033)	-0.0368***	(0.013)
uentry_sq	0.00431***	(0.0011)	-0.00510***	(0.0014)		
age			0.120	(0.10)		
age_sq			-0.00243*	(0.0015)		
preex12cum_sq			0.159*	(0.082)		
preex12cum_sq_a2539			-0.112*	(0.066)		
state2			-0.304*	(0.18)	-0.575**	(0.26)
state3			-0.417**	(0.18)	-0.352	(0.22)
state4			-0.103	(0.15)	-0.448**	(0.22)
state5			-0.446**	(0.19)	-0.150	(0.21)
logearn_sq					0.0538**	(0.026)
frmsize4					0.376	(0.23)
frmsize4_uentry					0.00685	(0.022)
Constant	-3.083***	(0.29)	-4.104**	(1.79)	-3.394***	(0.56)
Observations	4713		3071		1724	
PseudoR^2	0.0950		0.117		0.162	

Table B.9: Balancing Tests Women Practice Firm

Stratum	Degree of Polynominal	P-values >.10	>.05	>.01	Regressors
1	3	10	11	13	13
1	4	12	12	13	13
2	3	9	9	10	10
2	4	9	10	10	10
3	3	8	10	13	13
3	4	7	11	13	13

Table B.11: Balancing Tests Women SPST

Stratum	Degree of Polynominal	P-values >.10	>.05	>.01	Regressors
1	3	23	24	26	27
1	4	23	24	27	27
2	3	20	20	21	21
2	4	19	21	21	21
3	3	29	29	32	32
3	4	30	31	32	32

Table B.13: Balancing Tests Women RT

Stratum	Degree of Polynominal	P-values >.10	>.05	>.01	Regressors
1	3	13	14	14	15
1	4	9	10	14	15
2	3	14	14	15	15
2	4	11	13	15	15
3	3	14	14	16	16
3	4	14	15	16	16

Table B.14: **Propensity Score Estimates Men Practice Firm**

COEFFICIENT	Stratum 1	
age	0.0515	(0.067)
age_sq	-0.000821	(0.00085)
BER3	-0.231	(0.21)
earn	-0.00313	(0.0050)
frmsize4	0.419***	(0.14)
state1	0.447**	(0.18)
state3	0.490***	(0.16)
state5	0.363**	(0.18)
uentry	0.0181	(0.030)
uentry_sq	-0.00119	(0.0013)
Constant	-3.367***	(1.28)
Observations	5598	
PseudoR^2	0.0663	

Table B.15: Balancing Tests Men Practice Firm

Stratum	Degree of Polynominal	P-values >.10	>.05	>.01	Regressors
1	3	9	9	9	10
1	4	7	8	10	10

Table B.16: **Propensity Score Estimates Men SPST**

COEFFICIENT	Stratum 1		Stratum 2		Stratum 3	
a2529	0.225	(0.15)				
a3034_4044	0.476***	(0.10)				
a3539	0.318**	(0.15)			0.254	(0.19)
a4549	0.307**	(0.12)	-0.334	(0.26)	0.0794	(0.22)
married	0.133*	(0.069)	-0.0324	(0.087)	-0.105	(0.097)
qual_l	-0.0892	(0.11)	0.0145	(0.11)	0.292	(0.22)
qual_h	0.470***	(0.12)	-0.104	(0.17)	0.438***	(0.16)
BER3	0.152	(0.096)	0.429*	(0.24)	0.0447	(0.22)
logearn	0.128	(0.093)	0.0582	(0.096)		
earncens	0.541	(0.43)	0.0931	(0.49)		
industry6	0.0362	(0.088)				
industry57	-0.208***	(0.078)				

frmsize2	0.163*	(0.088)			0.275**	(0.12)
frmsize3	0.148	(0.12)			0.154	(0.17)
frmsize4	0.356***	(0.11)			0.193	(0.15)
preex12	0.183	(0.12)	-0.136	(0.12)		
preex12cum_sq	0.0195	(0.026)			0.0434	(0.033)
preex12_a4055	0.124	(0.10)				
popdens	0.203***	(0.064)				
popdens_sq	-0.104**	(0.042)				
uentry	-0.0176	(0.016)	0.0954***	(0.026)		
uentry_sq	0.00145**	(0.00068)	-0.00260***	(0.00100)		
a3039			0.000506	(0.14)		
a4044			0.233	(0.15)	0.266	(0.20)
a5055			-0.674**	(0.27)	0.227	(0.19)
BER2			0.0651	(0.22)	-0.133	(0.21)
preex12_a4555			0.529**	(0.26)		
a3034					0.139	(0.18)
qual_l_a3544					-0.547*	(0.33)
qual_l_a4555					-0.291	(0.29)
preex12cum_sq_a4055					-0.123**	(0.061)
Constant	-3.019***	(0.39)	-2.551***	(0.42)	-1.603***	(0.27)
Observations	5758		2846		1525	
PseudoR^2	0.0807		0.0567		0.0353	

Table B.17: Balancing Tests Men SPST

Stratum	Degree of Polynominal	P-values >.10	>.05	>.01	Regressors
1	3	17	20	22	22
1	4	17	20	22	22
2	3	14	14	15	15
2	4	11	13	15	15
3	3	15	15	17	17
3	4	13	14	17	17

Table B.18: **Propensity Score Estimates Men RT**

COEFFICIENT	Stratum 1		Stratum 2		Stratum 3	
a2529	0.970***	(0.24)	0.843***	(0.22)		
a3034	1.034***	(0.24)	0.683***	(0.22)		
a3539	0.840***	(0.24)	0.684***	(0.22)		

a4044	0.651***	(0.25)	0.499**	(0.22)		
married	0.0561	(0.090)	0.0817	(0.11)	0.00675	(0.18)
qual_l	-0.174	(0.14)	-0.193	(0.27)	-0.179	(0.23)
qual_h	-0.604*	(0.35)			-0.244	(0.43)
logearn	0.0752	(0.092)			0.106	(0.15)
frmsize2	0.325***	(0.12)				
frmsize3	0.377**	(0.16)				
frmsize4	0.511***	(0.15)				
preex12	-0.303*	(0.17)			0.213	(0.22)
preex12cum	0.352***	(0.13)				
preex12cum_sq	0.124**	(0.056)	-0.0654*	(0.038)		
state1	0.0417	(0.13)			1.031***	(0.27)
state2	-0.0132	(0.14)			0.0906	(0.36)
state3	0.0538	(0.12)			0.716***	(0.27)
state5	-0.290**	(0.15)			0.380	(0.31)
popdens	0.250***	(0.089)				
popdens_sq	-0.192***	(0.066)				
uentry	0.00237	(0.0062)	0.0428	(0.029)		
qual_m			-0.260	(0.24)		
BER2			0.322	(0.41)		
BER3			0.140	(0.43)		
logearn_sq			0.0196	(0.019)		
uentry_sq			-0.00240*	(0.0013)		
age					-0.0502	(0.13)
age_sq					0.000270	(0.0017)
pretx1					0.530*	(0.30)
Constant	-3.085***	(0.44)	-2.761***	(0.56)	-1.474	(2.35)
Observations	4551		2125		1030	
PseudoR^2	0.0814		0.0509		0.118	

Table B.19: Balancing Tests Men RT

Stratum	Degree of Polynominal	P-values >.10	>.05	>.01	Regressors
1	3	17	18	21	21
1	4	19	21	21	21
2	3	11	13	13	13
2	4	11	11	13	13
3	3	10	11	12	12
3	4	7	9	12	12

B.1.4 Common Support

Figure B.1: Densities of Propensity Scores for Women

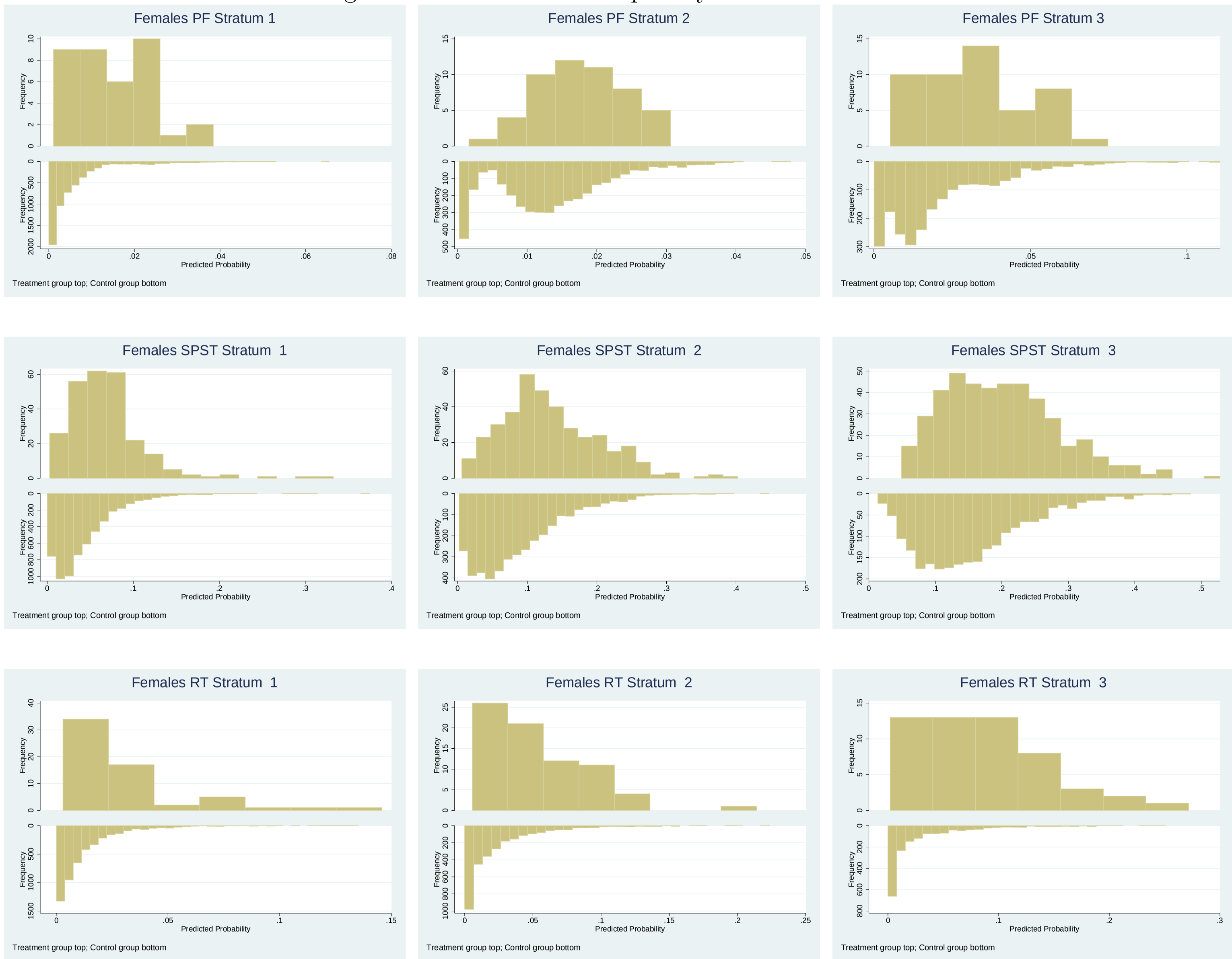


Figure B.2: Densities of Propensity Scores for Men

B.2 Background Information about the data

Other types of further training

In this study we are interested in active labor market programs for unemployed who have previously been employed and who have not already found a new job. However, we also want to give a short overview about other programs regulated by the labor promotion act (AFG) which we do not evaluate.

German Courses
The German Courses are intended for newly arrived immigrants. So the participants typically have not been employed in Germany before the German Course and hence are not part of the focus group of this study, the previously employed unemployed.

Career Advancement
These programs are typical programs directed at the employed, which were more important when the labor promotion act was introduced in 1969. By providing additional human capital the participant's risk of becoming unemployed should be lowered. Prime examples are courses in which the participants with a vocational training degree obtain additional certificates which allow them to independently run craftsman's establishments and to train trainees in the dual system of vocational training.

Wage subsidies
Wage subsidies are paid for the employment of formerly long-term unemployed and are intended to decrease the competitive disadvantage of these recruits for the period of familiarization with the skill requirement of the job. Even if the target group of wage subsidies are also unemployed we do not evaluate them because they require a job for which the wage subsidy is paid. This means provision of wage subsidies is already conditional on employment which is the success criteria for the other programs.

Any program which starts together with a job
For the same reasons why we do not evaluate wage subsidies we also do not evaluate any program which starts together with employment. Because we want to evaluate the program's effect on employment we do not consider programs which start together with employment.

Construction of the monthly panel

The IABS employment and LED benefit payment data are daily register data whereas the FuU training data gives monthly information about program participation. This study uses the merged data as described in Bender et al. (2005). From the merged data we construct a monthly panel. If the original daily data contain more than one spell overlapping a specific month we take the information from the spell with the largest overlap as the spell defining the monthly information.

The defining condition to be part of our inflow sample into unemployment is a transition from an employment month to a nonemployment month, in which the last employment month was between December 1992 and November 1993 and thus the first unemployment month was between January 1993 and December 1994. In order to divide nonemployment (to be precise: not employed subject to social security contributions) into unemployment and other states (like labor market leavers, transition into self employment, employment as civil servant) we additionally require a month with benefit payments from the employment office within the first twelve month of nonemployment or indication of participation in any labor market program in one of our data to be part of the inflow sample in unemployment.

Later on we aggregate the information further from monthly to quarterly information. Whereas the monthly employment information is binary the quarterly employment information can take the values 0, 1/3, 2/3 or 1.

We identify program participation if a person starts a program while being in the defining unemployment spell. The participant must not be employed in in the first month of the program. Otherwise we would consider such a program as a program which starts together with a job which we do not evaluate. In this case we would treat such a person as being employed. The exact identification of the program types will be explained in the following.

Identifying program participation

We identify participation in a further training program from a combination of FuU training data information, the benefit payment information and the employment status information. In principle, every participant in a further training program should be recorded in the FuU training data and we would not need the benefit payment data for identification of participation. There are two reasons to use the benfit payment data as well. First we find the training data to be incomplete, many recipients of training related benefits are not contained in the training data.[1] Only using the benefit payment data identifies these participants. Second, quite often the type of training in the training data is given very unspecific as "Other adjustment of working skills". The benefit payment data can give more information about these programs. Finally we need the employment status to identify participation because we only evaluate programs which start while being unemployed.

In the remaining part of this section we describe how we aggregate the benefit payment information and the training data information. The next section contains the exact coding plan. We disclose in detail which combination of information from benefit payment and training data we identify as PF, SPST or RT.[2]

Benefit payment information from the LED-data

The merged data we use contain three variables with benefit payment information from the original LED data, ("parallel original benefit information 1-3" [*Leistungsart im Original 1-3*] L1LA1, L2LA1, L3LA1). The main variable is L1LA1. If there are two parallel payment informations in the original data L1LA2 also contains information and only if there is a third parallel payment spell L3LA1 is also filled. In general we use L1LA1. Only if L1LA1 is not informative about program participation and L2LA1 is we use L2LA1 and only if L1LA1 and L2LA1 are not informative but L3LA1 we use L3LA1. The benefit payment information is given in time varying three-digit codes (for the coding plan see Bender et al. 2005). We extracted the program related information from the benefit payment information as given in table B.20. The main distinction regarding program participation is the distinction between no benefits at all or unemployment benefits/assistance on the one hand and program related maintenance benefits on the other hand. There are five types of program related benefits. Most important for us are the more general maintenance benefits while in further training and the more specific maintenance benefits while in retraining.

[1]Remember the purpose of the training data was only internal documentation. This might explain its incompleteness.

[2]More details about the benefit payment data and training data can be found in Speckesser (2004), Fitzenberger and Speckesser (2005) and Bender et al. (2005).

Table B.20: Aggregated types of benefit payment

German Abbreviation	Description
ALG	unemployment benefits
ALHi	unemployment assistance
UHG §41a	maintenance payment while in specific short term measure
UHG Fortbildung	maintenance payment while in further training
UHG Umschulung	maintenance payment while in retraining
UHG Darlehen	maintenance payment as a loan
UHG Deutsch	maintenance payment while in a German course

The original benefit payment information is given in three variables L1LA1, L2LA1 and L3LA1 with time varying three-digit codes.

Type of training from FuU-data

In this evaluation study one of the most important advantages compared to survey data is the information about the precise type of training. It allows us to identify homogeneous treatments for the evaluation. In the merging process, up to two parallel FuU-spells were merged to one spell of the IABS data because in many cases the FuU-data provided more than one parallel spell. These two parallel spells provide two variables indicating the type of course (*Maßnahmeart* [FMASART1, FMASART2]).

Aggregating the training type information Since type of treatment (*Maßnahmeart*) is often coded as "other adjustment"
(FMASART1=12 [*Sonstige Anpassungen*]) in the FuU-data, we increase the precision of information about the type of treatment by relying on the second parallel information about the type of training: The second FuU-spell is used if the first FuU-spell is coded as "other adjustment" (*"Sonstige Anpassungen"*) and a second spell includes a code different from 12. Such combined information of FMASART1 and FMASART2 is referred to as FMASART* in the following.

Combining the information

When using information from different sources, the sources may give differing information. If the training data indicated training participation and the benefit payment data did not or vice versa we relied on the source which indicated training for the following reasons. If somebody receives training related benefits it is more likely that the employment agency forgot to fill in the training data record than the agency wrongly induced payment of benefits. And if somebody is contained in the training data but does not receive maintenance benefits he either receives no benefits, which is possible while being in training, or receives unemployment benefits/assistance and the payment is just wrongly labelled.

If both training and benefit payment data indicate program participation but differ in the type of program we generally use the training data information. An example: the benefit payment indicates maintenance payments for further training and the training data indicates Retraining. We use Retraining from the training data. The only exception is unspecific program information from the training data "other adjustment". If in such cases the benefit payment data give specific information like Retraining we use the information from benefit payment data. All possible combinations of training and benefit payment information which we use to identifiy participation in one of the three programs are given in the following section.

Coding plan for the treatment information

This section gives the exact coding plans for identification of Practice Firm, SPST and Retraining. In general we identify program participation as start of a program in an unemployment spell before another employment begins. This means that we only identify a start of a program if the employment status in the first month of the program indicates no employment (BTYP$\neq$1).

Practice Firm

Practice Firm is a consolidation of the program types Practice enterprise and Practice studio from the FuU training data. There is no specific benefit payment type related to Practice Firms, rather the participants shall receive the general maintenance payment for further training. Since the training data are more reliable than the benefit payment data regarding type of the program we identify Practice Firm whenever FMASART shows the codes 11 or 12 independently of the payment information.

Program code	Label	Label in German
10	Practice enterprise	Übungsfirma
11	Practice studio	Übungswerkstatt

In table B.21 we show how often which combination of benefit payment information and program type information identifies *Practice Firm* in the two inflow samples.

Table B.21: Identification of *Practice Firm* with program type and benefit payment type: Frequencies

Program	Type of payment no benefits	UB/UA	maintenance benefits for short term training	further training	retraining	Total
Practice enterprise	0	0	0	106	0	106
Practice studio	0	0	0	110	2	112
Total	0	0	0	216	2	218

Women and Men together. BTYP$\neq$1 as an additional requirement.

Provision of specific professional skills and techniques

We identify SPST in the following cases.

(a) Identification from training data and benefit payment data
We identify SPST if the training data indicates the general program "Other adjustment" and the benefit payment information is no benefit payments, unemployment benefits, unemployment assistance or maintenance payments while in retraining.

Program code	Label	Label in German
12	Other adjustment of working skills	sonst. Anpassung der berufl. Kenntnisse

(b) Reliance on benefit payment data
We identify SPST if the program information from the training data is missing and the benefit payment information is maintenance payments while in further training.

Program code	Label	Label in German
-9	missing	fehlende Angabe

(c) Additional program from training data
We also identify SPST when another program of little quantitative importance but SPST–comparable content is recorded in the training data independent of the benefit payment information.

Program code	Label	Label in German
31	Further education of trainers and multidisciplinary qualification	Heran-/Fortbildung v. Ausbildungskräften/ berufsfeldübergreifende Qualifikation

(d) Additional combinatioin

Finally we identify SPST if the training data indicate the unspecific "other career advancement" and the benefit payment information indicates further training.

Program code	Label	Label in German
28	Other promotion	sonstiger Aufstieg (< 97)

In table B.22 we show how often which combination of benefit payment information and program type information identifies *SPST* in the two inflow samples.

Table B.22: Identification of *SPST* with program type and benefit payment type: Frequencies

Program	Type of payment no benefits	UB/UA	maintenance benefits for further training	Total
missing	0	0	549	549
Other adjustment of working skills	6	10	1158	1174
Other promotion	0	0	6	6
Further education of trainers and multidisciplinary qualification	0	0	9	9
Total	6	10	1722	1738

Women and Men together. BTYP$\neq$1 as an additional requirement.

Retraining

Retraining or longer "Qualification for the first labor market via the education system" is taking part in a new vocational training and obtaining a new vocational training degree according to the German dual education system. Additionally, but quantitatively of little importance we see the make up of a missed examination "Certification" as comparable to retraining because the result is the same. Furthermore and also only of marginal importance we see participation in the programs "Technican" or "Master of Business administration (not comparable to an american style

MBA)" while not receiving maintenance benefits as a loan as Retraining. Conventionally these two programs are considered as career advancement programs which we do not evaluate. Benefits as a loan would underline their character as career advancements.

(a) Identification from training data

We identify the following two programs as Retraining independent of the benefit payment information.

Program code	Label	Label in German
29	Certification	berufl. Abschlussprüfung
32	Retraining	Umschulung

(b) Reliance on benefit payment data

If the training data is uninformative and maintenance benefits for Retraining are paid we identify Retraining.

Program code	Label	Label in German
-9	missing	fehlende Angabe
12	Other adjustment of working skills	sonst. Anpassung der berufl. Kenntnisse

(c) Other programs from training data

Two other programs are identified from the training data. They typically also take two years full time and require an existing vocational training degree, hence are somewhat comparable to retraining in a narrower definition. Not identified if maintenance benefits are paid as a loan.

Program code	Label	Label in German
26	Technician	Techniker (<97)
27	Master of business administration	Betriebswirt (<97)

In table B.23 we show how often which combination of benefit payment information and program type information identifies *Retraining* in the two inflow samples.

Table B.23: Identification of *Retraining* with program type and benefit payment type: Frequencies

Program	Type of payment					Total
	no benefits	UB/UA	maintenance benefits			
			further training	retraining	loan	
missing	0	0	0	55	0	55
Other adjustment of working skills	0	0	0	13	0	13
Technician	0	0	0	0	0	0
Master of business administration	0	0	0	0	0	0
Certification	0	0	1	0	0	1
Retraining	2	2	219	137	0	360
Total	2	2	220	205	0	429

Women and Men together. BTYP$\neq$1 as an additional requirement.

Sample construction in comparison to LMW

Table B.24: Overview sample construction

	this paper	LMW
Inflow sample	starts unemployment spell in 93/94	starts unemployment spell in 93/94
Treatment group	starts a program within 24 months after beginning of unemployment spell	starts a program between beginning of unemployment spell and the end of 94
Control group	dynamic control group: does not start a program in the stratum of unemployment under consideration	static control group: does not start a program between beginning of unemployment spell and the end of 94
Treatment identification	training spell in the training participation data or income maintenance spell in the benefit payment data indicating program participation	training spell in the training participation data
Age restriction	25–55 years (25–50 in case of RT) in the year of entry into unemployment	20–53 years in the year of the (simulated) program start
Benefit payment restriction	controls have to receive unemployment benefits at least once during the first 12 months of their unemployment spell	recipience of benefits in the month before program start for participants and in the month before as well as in the month of the simulated program start for controls
Other restrictions	without East Berlin	last employment before defining unemployment spell not as trainee, home worker, apprentice, or in part-time with less than half of the usual ours; no foreigners
Sample size RT	Women: 189 (=61+75+53), Men: 219 (=107+79+33), numbers in parenthesis differentiated by strata	Women: 190, Men: 255
Sample size SPST (this paper) and short and long training (LMW)	Women: 1063 (=254+374+435), Men: 485 (=200+141+144)	Women: 557 (=209 (short) + 348 (long)), Men: 302 (=112+190)
Sample size non-participants	Women: 4585, Men: 5076 (not directly comparable to the dynamic control groups used in the paper)	Women: 2914, Men: 1690

B.3 Heterogeneous Treatment Effects by Target Profession

B.3.1 Retraining for Men

In this section we show heterogeneous treatment effects of retraining on men. We contrast retraining (RT) with target profession in construction with RT with other target profession (non-construction).

The effects are estimated in the same way as the non-disaggregated effects in the paper. We used the same propensity score specifications and bandwidth as in the paper.

Table B.25: Sample sizes: Retraining for men by target profession

	Stratum 1	Stratum 2	Stratum 3
Construction	40	29	19
Non-construction	50	40	13
missing	17	10	1
Total	107	79	33

Remark: The participants are classified according to the field in which they are retrained. This information is only available from the training participation data and hence is missing if participation is identified from the benefit payment data.

Figure B.3: Retraining (RT) for Men stratum 1 – Construction (top) vs. Non-construction (bottom)

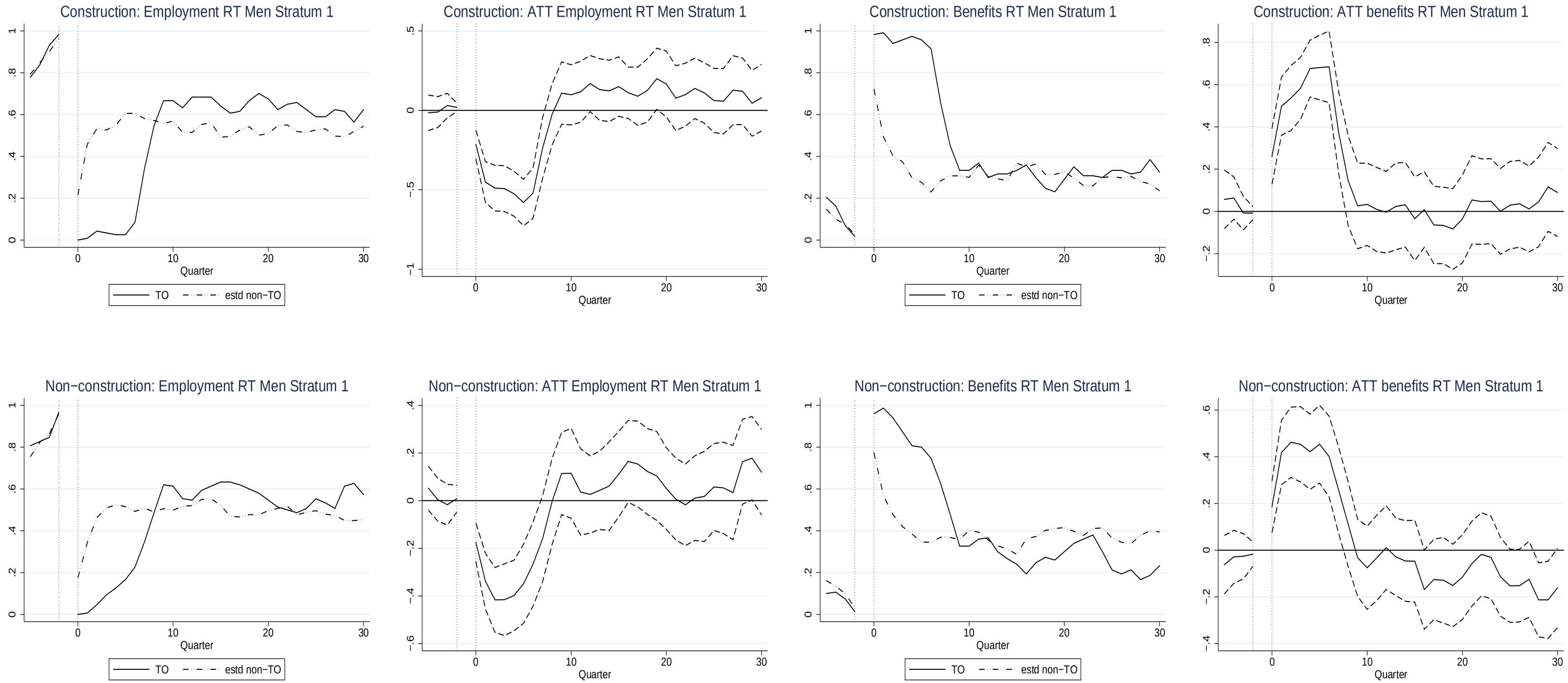


Figure B.4: Retraining (RT) for Men stratum 2 – Construction (top) vs. Non-construction (bottom)

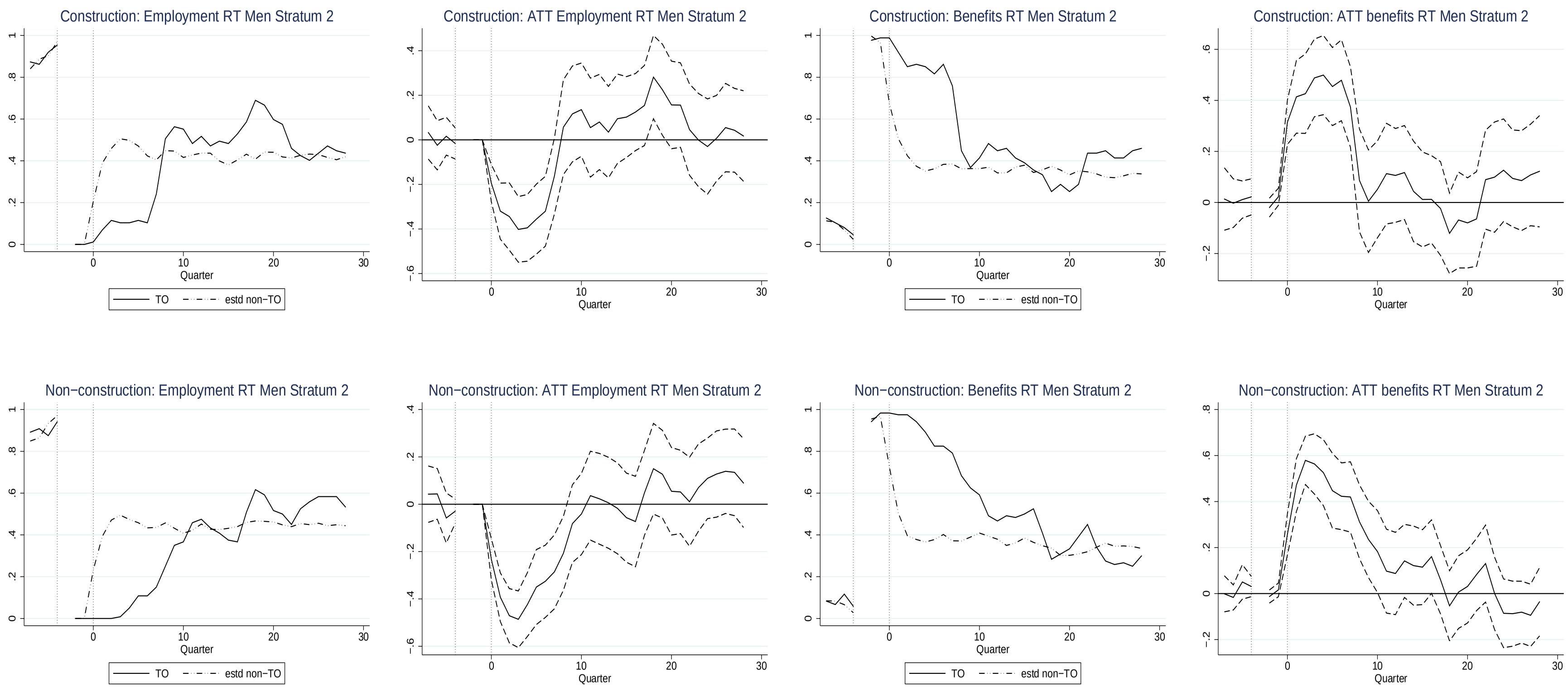


Figure B.5: Retraining (RT) for Men stratum 3 – Construction (top) vs. Non-construction (bottom)

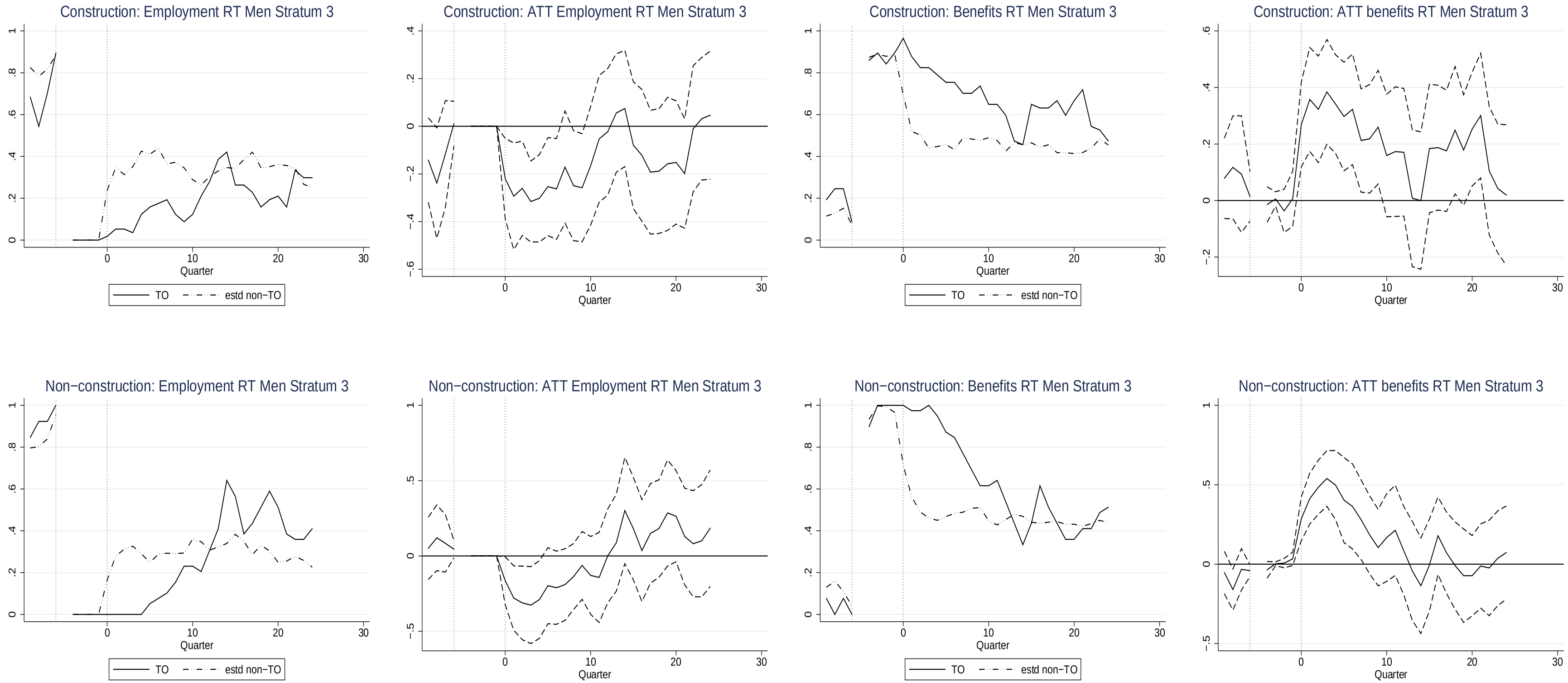


Table B.26: Cumulated and Average Employment Effects RT for Men – Construction vs. Non-construction

		Cumulated Effects			Average Effects	
Stratum	Target Profession	Q0–Q7	Q0–Q15	Q0–Q23	Q4–Q23	Q8–Q23
		Employment				
1	construction	-3.51 (0.44)***	-2.64 (0.91)***	-1.64 (1.37)	0.001 (0.064)	0.117 (0.071)*
1	non-construction	-2.52 (0.46)***	-2.02 (0.90)**	-1.42 (1.32)	-0.004 (0.061)	0.069 (0.066)
2	construction	-2.49 (0.47)***	-1.82 (0.91)**	-0.68 (1.34)	0.029 (0.063)	0.114 (0.071)
2	non-construction	-2.97 (0.41)***	-3.31 (0.91)***	-2.87 (1.39)**	-0.064 (0.064)	0.006 (0.070)
3	construction	-2.08 (0.56)***	-2.78 (1.15)**	-3.77 (1.81)**	-0.134 (0.084)	-0.106 (0.096)
3	non-construction	-1.97 (0.77)**	-1.87 (1.45)	-0.64 (2.34)	0.022 (0.106)	0.083 (0.112)
		Benefit Recipiency				
1	construction	4.29 (0.41)***	4.53 (0.88)***	4.43 (1.27)***	0.128 (0.061)**	0.009 (0.069)
1	non-construction	3.05 (0.46)***	2.91 (0.88)***	2.12 (1.24)*	0.030 (0.056)	-0.058 (0.061)
2	construction	3.45 (0.47)***	3.98 (0.80)***	3.83 (1.15)***	0.109 (0.054)**	0.024 (0.062)
2	non-construction	3.69 (0.42)***	4.98 (0.75)***	5.40 (1.11)***	0.176 (0.052)***	0.107 (0.057)*
3	construction	2.51 (0.54)***	3.68 (0.84)***	5.17 (1.36)***	0.192 (0.063)***	0.167 (0.077)**
3	non-construction	3.27 (0.67)***	3.84 (1.27)***	3.94 (1.84)**	0.111 (0.087)	0.042 (0.090)

Remark: The cumulated (average) effects are the sum (average) of the quarter specific average treatment effect on the treated over the respective quarters. $*$, $**$ and $***$ denote significance at the 10%, 5% and 1% level, respectively, and QiQj denotes quarter i to quarter j since beginning of treatment.

Lebenslauf

Persönliche Daten

Robert Helmut Alexander Völter
geboren 1975 in Berlin

Ausbildung

10/2000 - 9/2008	Promotion in Volkswirtschaftslehre Universität Mannheim Mitglied im Graduiertenkolleg Volkswirtschaftslehre und im Center for Doctoral Studies in Economics
07/2000	Diplom-Volkswirt
10/1995 - 07/2000	Studium der Volkswirtschaftslehre Rheinische Friedrich-Wilhelms-Universität Bonn
08/1998 - 06/1999	Ph.D. courses in Economics University of California, Berkeley
06/1994	Abitur
09/1989 - 06/1994	Heinrich-Hertz-Oberschule Berlin-Friedrichshain

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig angefertigt und die benutzten Hilfsmittel vollständig und deutlich angegeben habe.

Berlin, den 24. April 2008

Robert Völter