Do Reported End Dates of Treatments Matter for Evaluation Results? - An Investigation based on the German Integrated Employment Biographies Sample [1]

Marie Waller

Goethe University Frankfurt, CDSEM University of Mannheim, ZEW, f-waller@zew.de

Correspondence: Department of Economics, Goethe University, PO Box 11 19 32 (PF 247), 60054 Frankfurt am Main, Germany.

This version: November 24, 2006
Preliminary!

Abstract: This paper investigates the measurement error in end dates of further training programs in the German Integrated Employment Biographies Sample (IEBS) to gain insights on how measurement error in end dates of treatments affects evaluation results and on how to deal with this problem in future studies using the IEBS. Error-proneness of end dates in the IEBS is discussed, correction procedures are introduced and their influence on estimated employment effects is analysed using simple descriptive analysis, propensity score matching and a descriptive duration model. Though there is considerable measurement error in the end dates that can be corrected, the effect on evaluation results is modest, because for evaluation end dates are relevant only through indirect channels. The impact on causal effects based on matching is highest during the lock-in period and for long programs.

Keywords: evaluation, administrative data, measurement error, active labor market policies
JEL: C 14, J 68, H 43

---

[1] This study is part of the project "Employment effects of further training programs 2000-2002 – An evaluation based on register data provided by the Institute of Employment Research, IAB (Die Beschäftigungswirkungen der FbW-Maßnahmen 2000-2002 auf individueller Ebene – Eine Evaluation auf Basis der prozessproduzierten Daten des IAB)" (IAB project number 6-531.1A). The project is joint with the Swiss Institute for Internationale Economics and Applied Economic Research at the University of St. Gallen (SIAW) and the Institut für Arbeitsmarkt- und Berufsforschung (IAB). I gratefully acknowledge financial support by the IAB. All errors are my sole responsibility.

Contents

1 Introduction

A large and rich administrative data set, the Integrated Employment Biographies Sample (IEBS), became available recently. Covering about 1.4 million individuals and rich, daily information on employment, job search, transfer payments and active labor market programs, this data set is unique in Germany. It is the basis for the ongoing government conducted evaluation of recent years' labor market reforms and will probably be used for almost all empirical studies on labor market policies in Germany in the next years. The data are considered highly reliable, but end dates of further training programs are an exception to this. Mostly due to early drop-out, a considerable part of reported program end dates in the data is later than the end of actual participation. Because measurement error in end dates may influence standard evaluation results through several channels, it is difficult ex ante to predict how the problem affects results.

The IEBS has the advantage that due to its richness and its special structure it is possible to correct almost all relevant end dates. This advantage is exploited in this paper. Four different procedures to deal with the problem are introduced: a standard approach, a "naive" approach and two procedures using slightly different correction mechanisms. These four procedures are used to study through which channels and to what degree upward measurement error in end dates influences estimated employment effects. Descriptive employment rates, treatment effects using propensity score matching and a descriptive proportional hazard model are estimated using a framework with typical properties of evaluation studies like the probability of regular employment as the outcome, a focus on employment effects from the start of the program on and the consideration of program effects as opposed to pure threat effects. There are two aims of this exercise. The first is to gain knowledge on how to handle the problem in future studies using the IEBS. The second is to get insights on how strongly measurement error in end dates of treatments influences evaluation results in empirical studies in general. This might be helpful for studies using other administrative data sets, which are supposed to suffer from measurement errors in end dates that cannot be corrected. To the best of my knowledge, there is no guidance in the literature on this problem.

The remainder of the paper is structured as follows: section two discusses error-proneness of end dates in the data and possibilities for corrections. Section three presents the framework for the evaluation, discusses in what way end dates may influence results and introduces four procedures to handle the problem. Section four,

five and six study the impact of error-prone end dates on basic descriptive results, on matching results and on results using descriptive duration analysis, respectively. Section seven concludes.

2 Error-proneness and Possibilities to Correct End Dates of Further Training Programs in the IEBS

2.1 The Integrated Employment Biographies Sample

The Integrated Employment Biographies Sample (IEBS) is a new, rich administrative data set. It consists of a 2.2% random sample of individuals data drawn from the universe of data records collected in four different administrative processes.[2] The individuals in the IEBS are thus representative for the population made up by those who have data records in any of the four administrative processes. The data contains detailed daily information on employment subject to social security contributions, receipt of transfer payments during unemployment, job search, and participation in different programs of active labor market policy. Thus the IEBS is particularly useful to evaluate different parts of German active labor market policies in detail. It is the data set that is used for the evaluations of the so called *Hartz-Reformen*, several major labor market reforms of recent years. It is also used in other evaluation studies currently in progress.[3] Many others studies using the IEBS will certainly follow as the data set is unique in Germany concerning its largeness and richness in detailed information and as it will be updated in the future to always include recent years.

The IEBS collects information from four different administrative sources: the Employment History (*Beschäftigten-Historik*), the Benefit Recipient History (*Leistungsempfänger-Historik*), the Supply of Applicants (*Bewerberangebot*), and the Data Base of Program Participants (*Maßnahme-Teilnehmer-Gesamtdatenbank*).

The first data source, the Employment History, consists of social insurance register data for employees subject to contributions to the public social security system. It covers the time period from 1990 to 2004. The main feature of these data is detailed

[2] For detailed information on the IEBS see Hummel et al. (2005) and Bender et al. (2005). Our project uses a version of the IEBS that has been supplemented with additional information.

[3] See Biewen et. al. (2006) and Lechner and Wunsch (2006).

daily information on the employment status of each recorded individual. In evaluation studies this information can be used to account for the labor market history of individuals as well as to measure employment outcomes. For each employment spell, in addition to start and end dates, data from the Employment History contains information on personal as well as job and firm characteristics such as wage, industry or occupation.

The second data source, the Benefit Recipient History, includes daily spells of all unemployment benefit, unemployment assistance and subsistence allowance payments individuals in our sample received between January 1990 and June 2004. It also contains information on personal characteristics, on sanctions and periods of disqualification from benefit receipt. The Benefit Recipient History is important as it provides information on the periods in which individuals were out of employment and therefore not covered by the Employment History. In particular, the Benefit Recipient History includes information about the exact start and end dates of periods of transfer receipt.

The third data source included in the IEBS is the so-called Supply of Applicants, which contains diverse data on individuals searching for jobs. The Supply of Applicants data cover the period January 1997 to June 2004. The spells include detailed information concerning job search, regional information and personal characteristics, in particular about educational qualifications, nationality and marital status. They also provide information about whether the applicant wishes to change occupations, how many job proposals he or she already got and about health problems that might influence employment chances.

The fourth and final data source of the IEBS is the Data Base of Program Participants. This data base contains diverse information on participation in public sector sponsored labor market programs like training programs, employment subsidies and many more covering the period January 2000 to July 2004. This paper focuses on medium and long term further training (*Maßnahmen zur Förderung der beruflichen Weiterbildung*). Similar to the other sources, information comes in the form of spells indicating the start and end dates at the daily level, the type of the program as well as additional information on the program such as the planned end date, whether the participant entered the program with a delay, and whether the program was successfully completed.

2.2 Error-proneness of End Dates in the Program Data

The reliability of the data was checked very carefully within the project this study is part of. We ran extensive consistency checks of the records coming from the different sources, making use of additional information on the data generating process provided to us by the Institute for Employment Research. In addition, we consulted experts in local labor agencies and studied many examples in the data.[4] Concerning calendar dates, our conclusion is that start and end dates in the employment and benefit data are very highly reliable.[5] The reasons for this seem to be that start and end dates of employment spells are directly relevant for contribution rates and therefore indispensable for the administration. Information on the start and the end of transfer payments are, at the administrative level, directly linked to flows of benefit payments. Dates seem to be less reliable in program participation data and job search data.[6] A particular problem are the end dates of further training programs, a considerable part of which is not correct. But end dates of program participation are often necessary to define the participation in the program to be evaluated.

There are several reasons why end dates of program participation spells are error-prone. First, the correct reporting of end dates of program participation is not as important for administrative purposes as for instance of benefit spells, where the end date is directly relevant for the payment. Second, the end of program participation often changes after the date is first registered. This can be due to drop-out of the program, non attendance, change of course or shift of the course. If then the registered date is not corrected or if the correction does not reach the data set provided to the researcher, the end date of participation in the IEBS will be incorrect. Third, program end dates are registered by hand, which may cause mistakes especially because they often lie far in the future. Considering that program spells are almost never found too short in the data, but often too long, the third

[4] This work is documented in Bender et al. (2004, 2005).

[5] Concerning other aspects of the data, we came to the conclusion that the variable that indicates the status of the spell is very reliable in all four sources. Information in the employment data and the benefit data that is needed for administrative purposes is very reliable (for instance wage and transfer payments) but information not directly needed in the administration is less reliable in these sources. Fitzenberger, Osikominu and Völter (2006) discuss imputations of the education variable in the IABS, another data set including information of the Employment History. Personal information seems to be better in the job search and program participation data sources.

[6] For job search data it is possible to circumvent this problem by defining the labor market status using benefit and employment data.

problem seems to be of minor relevance, because typos would lead to errors in both directions.[7]

Given these reasons program end dates will always be a sensitive part of administrative evaluation data in Germany and countries with comparable administrative structures. There are two aspects determining the reliability. One is how and when information is registered during the administrative process itself. The other one is what rules the providers of the data use to define which piece of information of the administrative data bases will finally appear in the scientific data set as the program end date.[8] These rules have changed between the IEB versions 2 and 3 used in the studies currently available (2.05 is used in this study) and version 4 which is just becoming available. In the versions before version 4 the start date is taken from the *first record* [9], i.e. the information drawn when a participation spell first appears in the administrative data basis. The end date is taken from the *final record*, i.e. the information that is collected in the first draw occurring after the end date of the program. If for some reason there is no *final record* the end date is taken from the *first record* and thus equal to the so called planned end date (Hummel et al. (2005) and Kruppe and Oertel (2003)). Since version 4, the participation data come from the so called *Datenbank ISAAK V.1.03*. Here the end date is always taken from the information with the latest record date except this date lies later than one year after the program end. If the information on the end date in the record with the latest date is missing, it is taken from the record with the earliest date.[10] Thus a major difference between the rules seems to be that there is no distinction between *first record* and *final record* any more but the date of the record is important. Other differences probably occur due to the new ISAAK system itself. The change of procedure implies an improvement in the reliability of end dates for some cases, but new problems arise. A first look on version 4.0 has shown that on the one hand about 70% of the end dates corrected using the correction mechanisms suggested in this paper remain incorrect in version 4.0. On the other hand there have occurred new errors. Therefore it is not clear ex ante whether the new rules will reduce or

[7] Start dates are more reliable than end dates, probably because drop-outs are irrelevant and because they lie in the nearer future, so that fewer changes occur. In case of non-attendance start and end dates are per definition incorrect. In this case the correction of the end date leads to non-participation in a program and thereby also to a correction of the start date.

[8] Jaenichen et al. (2005) analyze some inconsistencies of the participation data that are related to the end date problem. One of their conclusions is that both aspects are relevant, but the problems in the registering of the data themselves might be the major problem.

[9] *First record* is *Zugangsdatensatz* and *final record Abgangsdatensatz* in German.

[10] (IEB User Manual V. 4, not published)

augment the error-proneness of the end dates. This paper studies implications of end date correction in version 2, which the studies currently available use, but the correction procedures might as well be used for version 4.

2.3 Possibilities for Corrections of Program End Dates

Concerning data checks and corrections, the IEBS has a great advantage: the fact that it includes four administrative data sources can be exploited to check plausibility and correct implausible information. It is thus possible to correct end dates and to analyze if and how errors in treatment end dates in real data lead to biased estimation results. This section discusses what information can be used for corrections in general, whereas section 3.3 explains how the corrections used in the study are implemented.

A constellation in the data set, that is evidently a contradiction, is a regular employment spell that starts before the end of program participation.[11][12] Imagine for instance a retraining program that starts after three months of unemployment. The spell continues for two years, a typical duration for a retraining program. But after four weeks, a regular employment spell starts in parallel to the participation spell. As dates in employment data are much more reliable than in participation data, the employment information indicates that the correct end date of program participation is after four weeks of attendance at the latest.

A second major possibility for corrections is provided by subsistence allowance spells. Subsistence allowance are payments of the labor agency to cover living costs of the participants of medium and long term further training programs. They are a subsidy to unemployment benefit or unemployment assistance for the time of the program.[13] With very few exceptions discussed later, all participants of medium and long term further training programs receive subsistence allowance for the complete time of the program (sometimes plus the weekend after the end of the program), a fact that proves true in the data. Dates of subsistence allowance spells are very reliable. Thus, if a subsistence allowance spell finishes before a program spell, one can conclude that

[11] To be eligible for further training, a person has to be unemployed. In theory, it is possible to be partly unemployed if one loses one of several jobs and is still registered as unemployed, but the number of these part time unemployed is almost zero.

[12] In this paper regular employment is defined as non-minor unsubsidized employment on the first labor market with a minimum length of two weeks.

[13] Short term training is not analyzed in this paper.

the end date of the program spell is wrong.

Third, in very rare cases, the living costs of participants are covered by an apprentice wage payed by a firm. The apprentice wage spell may be used in the same way as a subsistence allowance spell for corrections. Furthermore, among the additional variables in our version of the data, we have a variable indicating never attendance and a variable indicating the date if someone signed off before the regular end of a program. These two variables are often missing, but used with caution they can help to correct the end dates in some cases.

There is other information in the data which one might be tempted to use, but which would lead to a false correction in some cases. This is for instance the length of program spells. The law provides certain rules for the length of certain programs, but despite of this in practice there exist - though rarely - much longer programs. Therefore one should not change end dates in the data just because a spell is surprisingly long. While regular employment parallel to training programs is a contradiction, employment of a few hours only may occur in some cases and must not be used for the correction of program spells. Some participants receive aid from the European Social Fonds (ESF). This fonds allows flexible combinations of ESF programs, normal programs accompanied by ESF allowance and subsistence allowance. Therefore neither the ESF spells themselves nor subsistence allowance spells in connection with ESF spells are safe to use for corrections. For technical reasons it happens, though very rarely, that program participation is split into different spells in the data that can even overlap. Therefore different participation spells should better be connected instead of deleting part of them.[14] Information originating from the so called class data base are risky to use, because this information might be merged wrongly to the participant data.[15]

[14] For proposals how to deal with this problem see Jaenichen et al. (2005).

[15] Class data -as opposed to participant data - is on the courses themselves. It is merged to the participant data using a course identifier.

3 Procedures to Handle Error-prone End Dates in Evaluation

3.1 Treatments and Sample

This section introduces the framework of the analysis, discusses how measurement error in the end dates may influence evaluation results within this framework and introduces four different procedures to handle error-prone end dates. The focus of this paper is on two different treatments: further training and retraining. Further training (FT) is a medium length training program lasting usually several months. In classroom or in a practice firm, new professional skills are trained and existent knowledge is updated. Retraining (RT) leads to a new degree within the German vocational training system, it lasts typically two years. Other public sector sponsored programs, like short term training or wage subsidies, are not evaluated, but they are considered within the multiple framework by excluding individuals attending these programs from the control group.[16] In this study the effect of the program itself (as opposed to a threat effect) shall be evaluated and therefore programs are only counted if the unemployed has participated a minimal amount of days. The limit has been set considering program aims and the distribution of planned program durations to 28 days for further training and 181 days for retraining.

For the rest of this paper, the focus is on an inflow sample into unemployment consisting of individuals living in West Germany who became unemployed between the beginning of February 2000 and the end of January 2002 after having been continuously employed for at least three months. Entering unemployment is defined as quitting regular employment and subsequently being in contact with the labor agency (not necessarily immediately) either through benefit receipt, program participation or a job search spell.[17] In order to exclude individuals eligible for specific labor market programs for young people and individuals eligible for early

[16] In the relatively rare case where one person has several participation spells within one unemployment spell, the spells are connected if there are at most 14 days in between two spells. If this concerns two different programs (for instance short term training and retraining), the connected spell is assessed as the more important program (retraining in the example). If a person participated in several programs within one unemployment period with an interruption of more than two weeks, the first program is evaluated.

[17] Note that this implies that the same individual may appear more than once in the evaluation sample. Approximately ten percent of the individuals are represented by more than one unemployment spell according to the above definition.

retirement schemes, only persons aged between 25 and 53 years at the start of their unemployment spell are considered.

3.2 Relevance of Program End Dates for the Estimation of Employment Effects

There exist several studies on measurement error in the treatment variable. Molinari (2005) develops limits for treatment effects in the case that the treatment variable has missings in survey data. Battistin and Sianesi (2006) characterize the bias if treatment status is mismeasured and provide bounds. Lewbel (2004) develops GMM estimators for three scenarios: the probability of treatment is known or an instrument with three values or an instrument with two values that is conditional independent of the outcome is available. The problem analyzed in this paper is different in two respects. First, the problem itself is more complicated, because the measurement error in the end date may influence the results not only through the treatment indicator. But second, using the IEBS data it is possible to correct the end dates. Therefore the aim of this paper is to develop procedures to correct the end dates and then to analyze how and to what extent wrong end dates influence different results.

When evaluating the employment effects of training programs, upward measurement error in end dates of program participation may have an effect through different channels. Using descriptive employment rates or matching, program end dates have no direct effect on the results but may bias them indirectly through outcome measurement and through the treatment indicator. First, if the outcome is measured as regular employment or nonemployment (including every other status including program participation), too late end dates of programs lead to a contradiction: the researcher observes program spells and regular employment spells in parallel for some time. A decision whether to count this time as employment or program participation (and thus non-employment) is necessary and will influence employment rates and treatment effects. Second, end dates define the actual length of program participation, which can be relevant for the decision if a program has been attended long enough to be counted for evaluation. Too late end dates can lead to measurement error in the treatment indicator: it may indicate participation, although it should indicate non-participation, as in reality the participant did not attend long enough. Measurement error in the end dates influences the results more directly in estimation designs in which it is of importance if a participant is in a program

at a certain point in time and if a program has been completed or not. This is for example the case if attending an uncompleted program and having attended a program in the past are considered separately using duration analysis. In conclusion, there exist different channels through which measurement errors in program end dates may bias evaluation results, but the end date does not enter the estimation directly (for instance as a regressor) and it is therefore difficult to predict the direction and magnitude of a potential bias. Four different procedures how to deal with the error-prone end dates are discussed in the following.

3.3 Four Procedures to Deal with Error-prone End Dates

3.3.1 Procedure 1

The underlying idea of procedure 1 is that program participation is the most important information in a data set mainly created for evaluation studies. Therefore participation spells are taken as they are in the data. If a participation spell conflicts with a regular employment spell, the participation spell is given priority. This rule is implemented for the measurement of the outcome, but also for the measurement of the labor market status before the relevant unemployment period in case that a former program spell conflicts with a regular employment spell. Procedure 1 is called the "naive" procedure, because a close look at examples in the data reveals that end dates of employment spells are more reliable than end dates of program spells.

3.3.2 Procedure 2

In procedure 2 regular employment spells are given priority in case they conflict with program spells. The rationale is that employment dates in the IEBS are very reliable, because the length of the spells is directly relevant for pension payment. The rule to give priority to employment information is applied for measurement of the outcome as well as measurement of labor market status before the relevant unemployment period in case of conflict, that exists in those cases where the researcher is forced to take a decision. But note, that no ex ante correction of the program end dates implemented. This implies that for the decision, whether a program has been attended or attended long enough, the participation spells are taken as they

are in the original data.[18] Procedure 2 is called the standard procedure, because it seems to be the best choice if one does not want to implement an explicit correction mechanism, but is convinced of the reliability of employment data.

3.3.3 Procedure 3

Procedure 3 works as procedure 2, but in addition a mechanism to correct end dates of participation spells for further training programs is implemented at the beginning of the data preparation. This works as follows: the first step of the correction mechanism uses regular employment for correction. To define periods of regular employment, spells of non-minor unsubsidized employment with positive wage are connected and overlapping spells are consolidated. If regular employment starts before the end of a program participation spell, program spells are assigned the date of the start of the employment spell minus one day, which is the last possible day of program attendance, according to what we know from the data. Eventually this assigned date will be used to cut off the program spell.

The second step prepares the use of subsistence allowance spells for correction. Further training spells are marked, if there is no reason - like ESF or an apprenticeship spell - that the participant should not - for the whole program or not at all - have received subsistence allowance. In these cases correction using subsistence allowance spells would be too risky, because combinations of different payments are possible in these cases. Next, subsistence allowance spells are connected. After this, the programs that should be funded in theory and have subsistence allowance spells in the data that start about the same time as the program itself are assigned the end date of the connected subsistence allowance spell. This date indicates the last day the participant has been in the program according to the benefit information. As a third step, the same is done for the rare further training programs for which living costs are covered by apprenticeship wages.

From these three steps, program spells can be assigned at most two dates. The earlier one is chosen as the new program end date. The end date is replaced in the original data set.[19]

As a further correction, programs that have not been attended according to the additional variables “program success: nonattendance” or have a too early date

[18] Remember from section 3.1 that program spells with a gap shorter than 14 days are connected.

[19] No other changes in the data are used later.

in the variable "signed-off at" are recoded to no program. Because the reliability of these variables seems to be restricted, they are only used if there is no information on subsistence allowance, ESF or an apprenticeship spell, which is already an indication that the program has not been attended.

3.3.4 Procedure 4

Procedure 4 works as procedure 3, with the only difference that the first step in the correction mechanism - the ex ante correction due to regular employment - is not pursued. As in procedure 2 and 3, employment still dominates participation if they conflict for the measurement of outcome or the time before unemployment. As a result the difference between procedure 3 and procedure 4 is that in the latter, too short participation will not be assessed as non-participation if the program is too short only with an end date correction using employment spells.

On the one hand it can be argued that procedure 4 is better than procedure 3, because procedure 3 is biased in the following way: unsuccessful participants will never have their end date corrected by the ex ante correction using employment spells, because per definition they have no regular employment spells starting before the end of the program. If there exists drop-out of unsuccessful participants not registered in the data and not uncovered by other correction steps, unsuccessful candidates will be over-represented in the treatment group. This leads to a downward bias of the treatment effect. But on the other hand one might also argue that Procedure 3 is preferable, because is makes as many reliable corrections as possible with the IEBS data. Therefore both procedures are used in this paper.

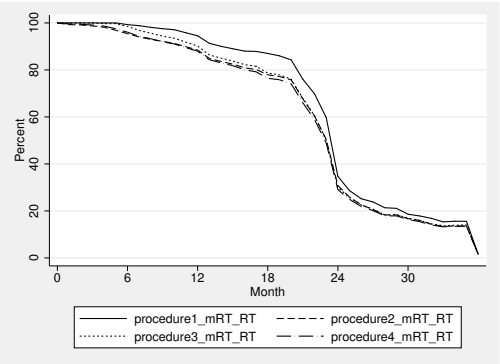
4 Impact on Basic Descriptive Analysis

4.1 Impact on Continuing Attendance on a Program

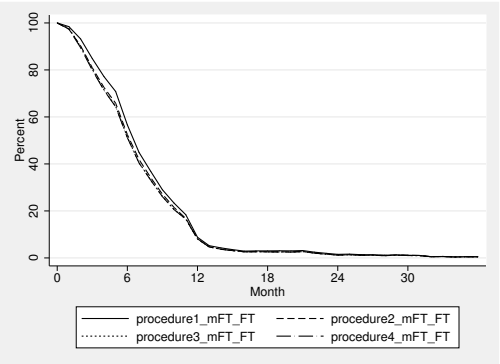
Figure 1 shows the rate of participants of RT (FT respectively) who are still attending an RT (FT respectively) program in the relevant month.[20] Month zero is the month in which the programs start, thus the participation rate is 100%. In

[20]Here, not the ex ante length of a program is shown, but the validation of the outcome. This implies that for procedure 2, 3 and 4 a month with a regular employment spell and a program spell is counted as a month in employment and not in a program.

Figure 1: Rates of Continuing Attendance on a Program

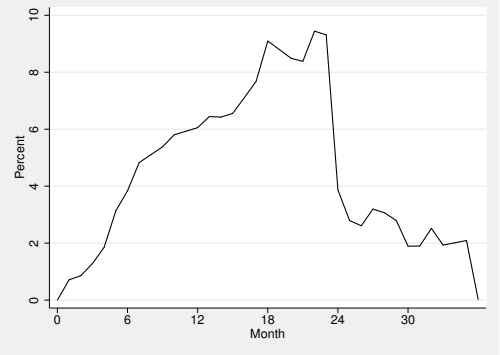


Participation Rate RT Participants

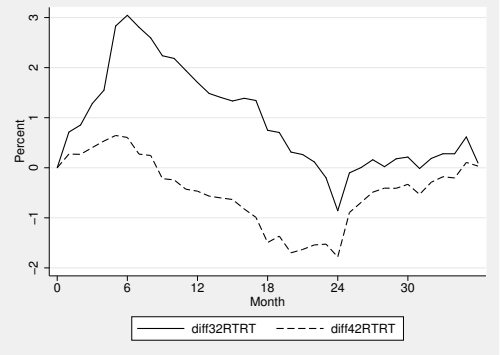


Participation Rate FT Participants

Figure 2: Differences of Continuing Attendance on a RT Program



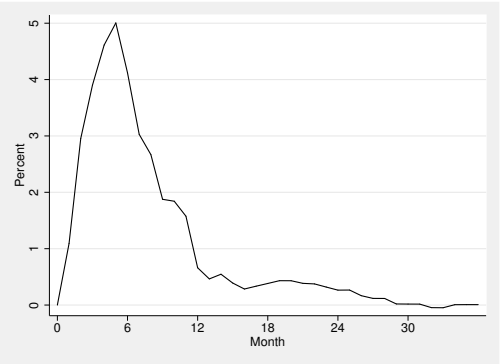
Procedure 1 minus Procedure 2



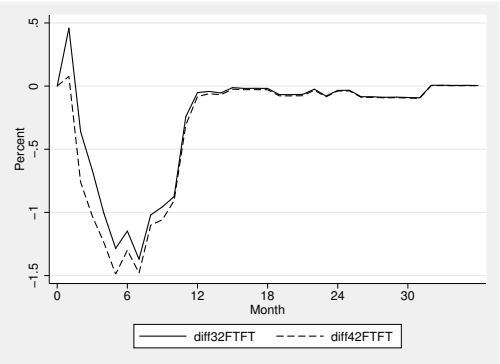
Procedure 3 minus Procedure 2 and Procedure 4 minus Procedure 2

each figure there is a graph for each of the four procedures. The graphs for RT show, that most participants stay in the program for 18 to 24 months, but some stay even 36 month in an RT program. About 20% of the participants leave their RT program within the first 18 month by dropping out. The individuals taking FT stay in the program much shorter, after 10 month the large majority has left the program. The differences between the procedures are larger for RT than for FT, because RT is the longer program and therefore the number of month where difference may occur is higher for RT, but the directions of the differences are the same. For RT at month 18 procedure 1 suggests the highest participation rate. This is because in procedure 1 program spells are counted, if they conflict with employment spells. Thus participants that are already in employment, but still have a (wrong) program spell, are counted as participants. Figures 2 and 3 are just another way of presenting the results, they show the differences of the graphs of figure 1. The highest difference between procedure 1 and 2 appears at the months before the end of the typical planned length of the programs, because at that time the most differ-

Figure 3: Differences of Continuing Attendance on a FT Program



Procedure 1 minus Procedure 2



Procedure 3 minus Procedure 2 and Procedure 4 minus Procedure 2

ences between reported end dates and true end dates occur. The participation rates of procedure 2, 3 and 4 are similar, here two effects compensate each other. On the one hand the corrections of procedure 3 and 4 prevent that drop-outs are still counted as participating after dropping out. This causes a lower participation rate for procedure 3 and 4. The effect is quite small because it is only relevant as long as the individuals are unemployed, because once they enter employment they are counted as employed and thus not as program participant using procedure 2, 3 or 4. On the other hand the corrections provoke that more RT program spells are shorter than six months (one month for FT) and therefore not counted, which leads to a non decreasing graph for the first six months (one month for FT) for procedure 3 and a slowly decreasing graph for procedure 4. The graph for procedure 2 decreases from the beginning on, because there are programs that are not corrected ex ante and therefore valid, but according to the rules of procedure 2 once an employment spell starts, this is counted as employment (and no program participation any more). In procedure 4 some program spells shorter than six months occur, because there is no ex ante correction using employment spells. These short programs not being in the sample, the program participation rate is higher for procedure 3 (and to a lower extend for procedure 4) than for procedure 2 in the beginning.

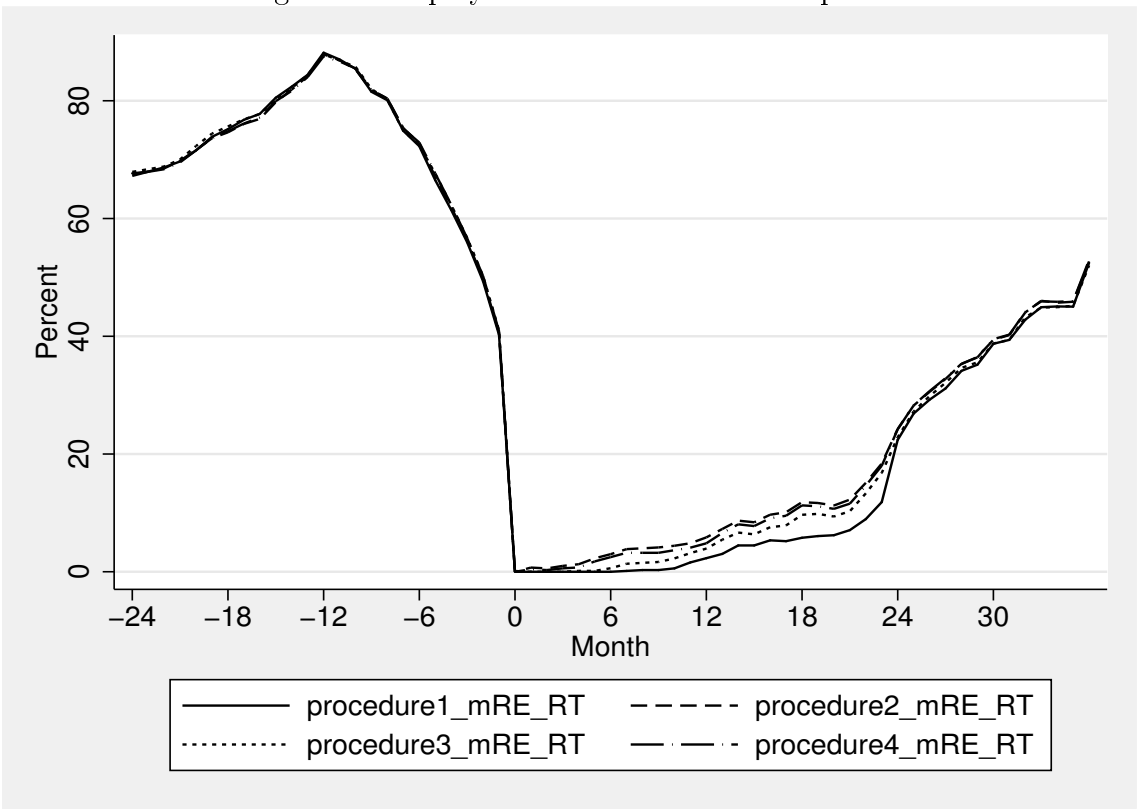
4.2 Impact on Employment Rates of Participants

4.2.1 Graphical Evidence

In the following the impact of the different procedures on the employment rates of participants is discussed, because the channels how the different procedures influence

results are analyzed easiest in studying descriptive evidence. Figures 4 and 5 show the employment rate of RT and FT participants respectively for each month before and after treatment start (month zero) for each of the four procedures.[21] Figures 6 and 7 are again just another way of presenting the results by showing the differences of the graphs of figures 4 and 5.

Figure 4: Employment Rate of RT Participants



Procedure 1 underestimates the employment rate up to six percent points as compared to procedure 2. This is because when measuring the outcome, program participation spells are given priority to regular employment spells. Figures 6 and 7 show that the magnitude of the effect is almost the same for FT and RT, but the period where differences occur - when program spells are wrong due to early drop-out of some participants - is longer for RT participants.

Using a correction mechanism results in a slightly smaller employment rate than the standard procedure (procedure 2). The measurement of the outcome is the same for procedure 2, 3 and 4, thus the difference in the employment rate must be due to differences in the validation of programs. In procedure 4 program spells are corrected (using mainly subsistence allowance spells) and therefore more non-attenders and

[21] For later months participation rates might be a little underestimated, because employment data of year 2004 is not yet complete.

Figure 5: Employment Rate of FT Participants

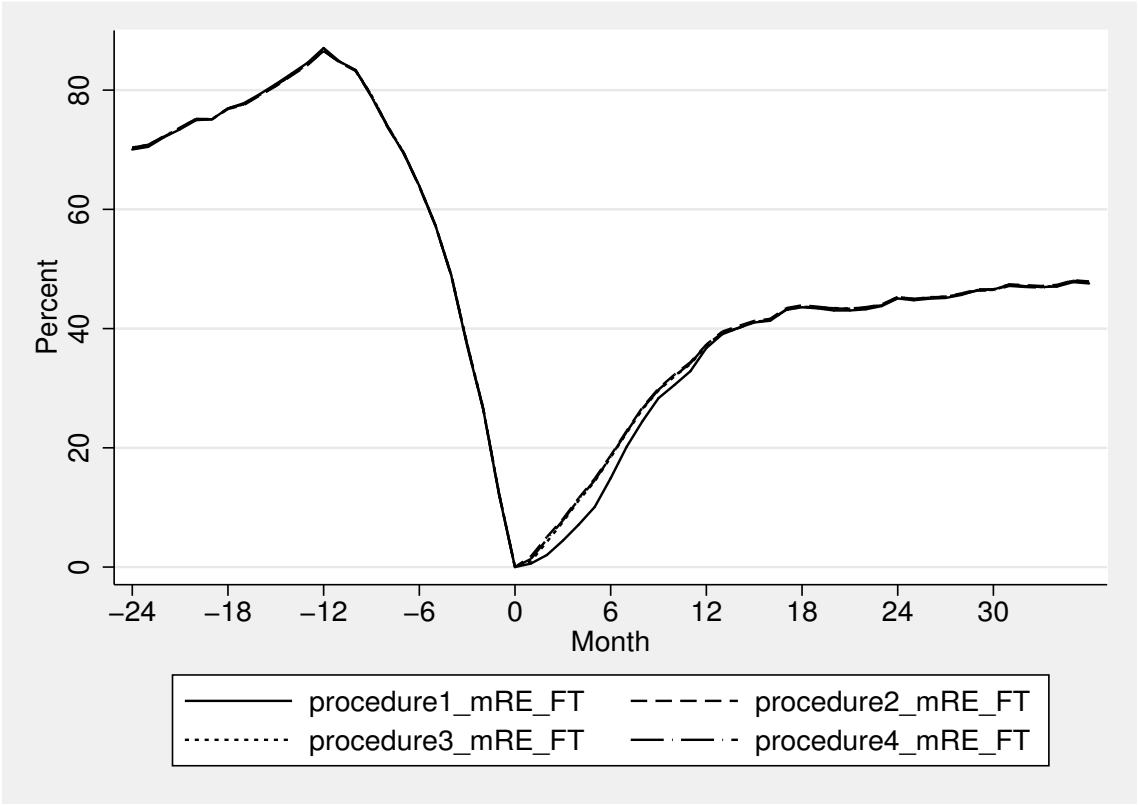
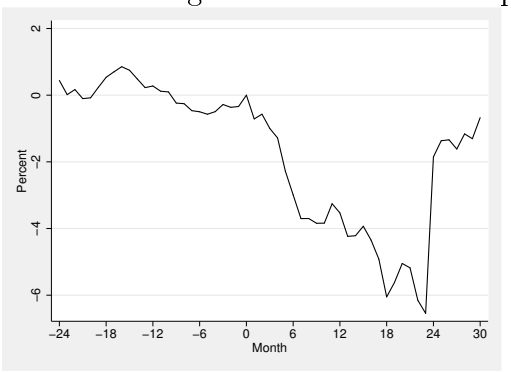
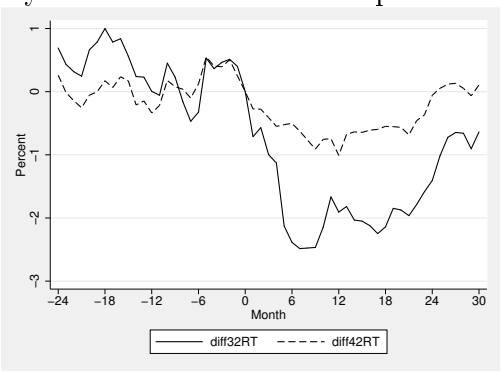


Figure 6: Differences Employment Rate of RT Participants



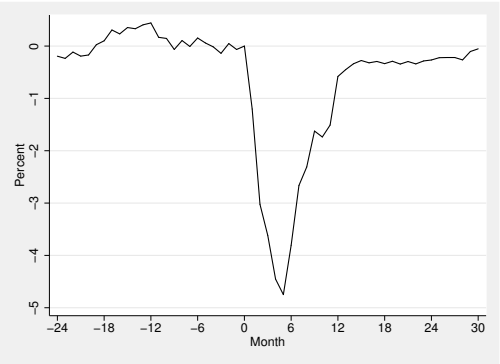
Procedure 1 minus Procedure 2



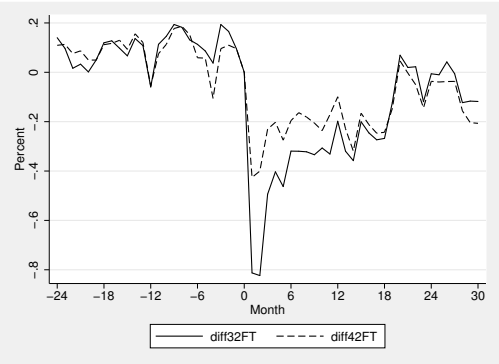
Procedure 3 minus Procedure 2 and Procedure 4 minus Procedure 2

very early drop-outs are not counted as participants. As figure 6 shows, this leads to a lower employment rate of participants, which indicates that non-attenders and very early drop-outs on average have higher employment rates. The employment rate using procedure 3 is even lower, which must be the case, because the difference between procedure 3 and 4 is that procedure 4 does not use employment spells for ex ante correction. Those spells that are corrected due to starting regular employment spells and are not in the treatment group for this reason have a higher employment

Figure 7: Differences Employment Rate of FT Participants



Procedure 1 minus Procedure 2



Procedure 3 minus Procedure 2 and Procedure 4 minus Procedure 2

rate per definition. Figures 6 and 7 show the differences in detail. According to Procedure 2 the employment rate of RT participants is up to 1 percent point higher than the employment rate according to procedure 4 and two percent points higher than the employment rate according to procedure 3. These differences are smaller for FT participants. Differences between procedure 2 and 4 last longer than differences to procedure 1, because the former are due to selection effects and the latter are due to outcome measurement.

4.2.2 Details on the Impact of the Corrections

This section shows in detail how the different procedures influence results within the given framework. The intention is to give a better understanding of how the procedures work and how measurement error in end dates may bias results. There are 2631 valid treatments using procedure 2. The end dates of only 1,9% of these treatments may not be checked with a correction procedure. Thus, it is possible to check almost all relevant participation spells. Only for 50 relevant spells the end date may neither be confirmed nor corrected. Some of these are programs with ESF for which a check and, if relevant, correction would be possible but does not seem safe enough. Some of these 50 programs were probably not attended, but this cannot be decided for sure and therefore the original data are kept as they are. It might also be the case that the subsistence allowance spell is missing in the final data set. In sum, the correction procedure leaves almost no open cases.

Table 1 gives for each procedure the number of valid unemployment spells, valid FT and RT treatments and the duration of the corrected and consolidated program spells. There are less valid employment spells using procedure 1 due the condition

of entering unemployment out of three months of employment. This condition is met a little less often in procedure 1, because also participation in earlier programs dominates earlier employment spells. Less programs are valid in procedure 3 than in procedure 4 and even less in procedure 2, because the more corrections are done, the more program spells are affected by the minimum attendance criterion. If the duration of the program spells are compared considering only those, that are valid in every procedure, obviously the average length is shorter the more corrections are pursued, irrespective if the length of the consolidated spell itself or the length of the spell until the beginning of a regular employment spell is considered. Considering the average length of those programs valid in the respective procedure, but not necessarily in all procedures, sample differences make this picture less clear.

Table 1: Programs in the Different Procedures

Procedure	1	2	3	4
Valid unemployment spells	51840	52539	52528	52539
Valid FT treatments	1918	1948	1928	1935
Valid RT treatments	673	683	642	664
Average duration FT	216.61	216.83	204.82	206.79
... for programs always valid	217.01	215.16	203.82	205.39
... and until employment only	217.01	206.63	203.02	203.02
Average duration RT	733.92	724.69	699.18	706.67
... for programs always valid	738.67	735.51	701.09	713.03
... and until employment only	738.67	713.01	698.17	698.94

Program duration is the ex ante program length of the consolidated program spell. This is different from figures 1 to 3 where the outcome in the relevant month is shown and thus for procedures 2, 3 and 4 employment spells that dominate program spells when measuring the outcome are already considered. Thus for figures 1 to 3 the length of the length of the program spell until employment is relevant.

Concerning the employment rates, first consider the differences between procedure 1 and procedure 2. The major part of the difference is due to the different handling of contradictions when measuring the outcome as explained above. This clearly leads to a downward bias in employment rates and treatment effects for procedure 1 as compared to procedure 2, where regular employment always dominates program spells. For 7.8% of the treatments valid in all procedures regular employment starts on average five months before the end of the original program spell. In addition to this, new programs and contradicting employment spells may start later on.[22]

[22]Minor differences between procedure 1 and 2 can arise from differences in the sample of valid

Second, consider the differences between the procedure 2 (the standard procedure) and procedure 4 (with corrections but without ex ante cutting off program spells due to employment spells), which may be explained through sample differences only, because the measurement of the outcome is the same. Some program spells which are wrongly classified as long enough without corrections, are too short to be evaluated or not attended when using corrections. In other words, the treatment indicator will in some cases indicate participation using procedure 2 and nonparticipation using procedure 4. The descriptive evidence suggests that the employment rate is lower with corrections, so these "drop-outs" or "non-attenders" seem to be more successful on average. If they are also more successful when controlling for selection, the treatment effect will be upward biased without corrections. But the sample differences are very small. While end dates change quite often due to corrections (out of the 2589 valid treatments in procedure 2 and 4, 12.4% (322) have an earlier end date due to correction) and the corrections are often quite severe (on average 103 days, 52 days is the median, 10% have corrections less than 2 days and 5% more than 407 days), only very few corrections influence the sample and can thus influence the employment rates.[23] Due to the corrections, 42 treatments valid in procedure 2 are not valid in procedure 4 (31 due to correction based on subsistence allowance, 9 due to an early signing-off date before the regular end of a program and 1 due to the indicated non-attendance). In conclusion, this shows that a considerable amount of end dates is corrected, but this correction has few implications, because the end dates do not directly influence the results. Only very few corrections have an indirect influence through sample changes because of the minimal length criterion.

Third consider Procedure 3, which involves an additional ex ante correction using employment spells compared to procedure 4. Due to this 1,16% of the treatments in procedure 4 are not valid in procedure 3.[24] These are per definition treatments that lead to employment. Thus employment rates and treatment effects estimated using procedure 3 will be slightly lower compared to procedure 4.

---

unemployment spells. If there is a valid unemployment spell in both procedures, differences in the validity of programs may evolve because of different spell consolidation, but they are rare (16 programs are valid in procedure 1 and not in 2 and 7 programs are valid in procedure 2 but not in 1).

[23] The overall sum of corrections in the data is of course much higher. Here only those programs relevant as treatments in the framework of this study are considered.

[24] Of those 2570 programs that are valid in procedure 3 and 4, 3,27% are further corrected in procedure 3, on average 140 days (median 62 days).

5 Impact on Treatment Effects Using Matching

5.1 The Method

In this section the impact of the four procedures on treatment effects using matching methods is studied. The matching approach used here is the one used in Biewen et al. (2006) and Fitzenberger et al. (2006) use. The approach uses a multiple framework as proposed by Lechner (2001) which in the case of this paper allows to estimate the effect of FT and RT separately against nonparticipation (defined here as participating neither in FT nor in RT nor in any other public sector sponsored program). Furthermore, the work of Frederikson and Johansson (2003, 2004) is taken into account. These authors show that a purely static matching approach leads to a biased estimator in settings where participants may start a program at different times during their unemployment spell and if program participation is relatively frequent. Therefore, building on the ideas of Sianesi (2003, 2004) Biewen et al. (2006) estimate treatment effects conditional on the elapsed unemployment duration at program start. They distinguish between treatments starting during months 0 to 3 of the unemployment spell (stratum 1), treatments starting during months 4 to 6 (stratum 2) and treatments starting during months 7 to 12 (stratum 3). In each of these strata, they define individuals to be undergoing treatment if they start the program under consideration during the time period defined by the stratum. Individuals not starting any program during the time window in question are in a "waiting" state because they are not treated at this point but may be treated later. Biewen et al. (2006) then carry out the evaluation for each stratum separately, circumventing the problem discussed above. A dynamic average treatment effect on the treated is estimated. It is interpreted as the effect a program has if it is taken in a certain stratum in contrast to taking no program in this stratum, under the condition that a person is still unemployed at the beginning of this stratum. For details compare Biewen et al (2006). Here, to study the implications of error-prone end dates, the treatment effects for women in West Germany attending an FT or RT program in the first three month of their unemployment spell (stratum 1) are discussed. The evaluation starts at the beginning of the program, because this is when programs start to have effects on the probability of employment of participants and because the end of the program may be endogenous.

To estimate the counterfactual outcome, matching based on the propensity score is used. The counterfactual is estimated by local linear matching on the propensity score and the calendar month of the start of the unemployment spell. As the kernel

function in the local linear regression a product kernel is used. Because the relevant control groups are very large, matching on the calendar month is exact, while the bandwidths for the propensity score are calculated by the Silverman rule of thumb.[25] The propensity scores are estimated using a probit. A large variety of characteristics of person, family, region and last job as well as health status, proxies for motivation, employment history and the number of job proposals are considered. For each estimation a separate specification is chosen according to the significance of variables and the balancing test of Smith and Todd (2005). The standard errors are computed using bootstrapping.[26]

5.2 Impact of the Different Procedures on Employment Effects

To see how the different procedures influence the treatment effects using the matching approach presented above, figures 8 to 13 show the average treatment effect on the treated for women in West Germany participating in FT or RT respectively during the first three months of their unemployment spell versus nonparticipation in a public sector sponsored program at that time. The outcome variable is the probability of regular employment at the respective month after the start of the program. The solid lines represent the treatment effects for different procedures. The months in which all persons are unemployed are not shown. Month zero is the start of the program. The dashed lines show the 95% confidence interval. Note that the confidence intervals are valid for the check whether the treatment effect is significantly different from zero and not whether the procedures differ.

Figures 8 and 9 show that, during the time when many participants are in the program and the treatment effects are negative (lock-in-effect), treatment effects both for FT and for RT differ remarkably if one uses procedure 1 instead of procedure 2. For FT, Procedure 1 overestimates the lock-in-effect about 5.28 percentage points at month 6 where the difference reaches its maximum and the average treatment effect on the treated (ATT) is -17.23% (10.60 percentage points for RT in month 19 at an ATT of -30.80). In the positive area of the treatment effects, when participants have finished the program, there are almost no differences for the two procedures

[25] The estimation procedures used in this section have been implemented in Stata by Aderonke Osikominu. Many thanks to her for the permission to use them for this study.

[26] For the estimation details see Biewen et al (2006). They use a leave-one-out crossvalidation procedure for the bandwith choice, which does not seem necessary for the purpose of this study.

Figure 8: Treatment Effect for FT Participants with Procedure 1 and Procedure 2

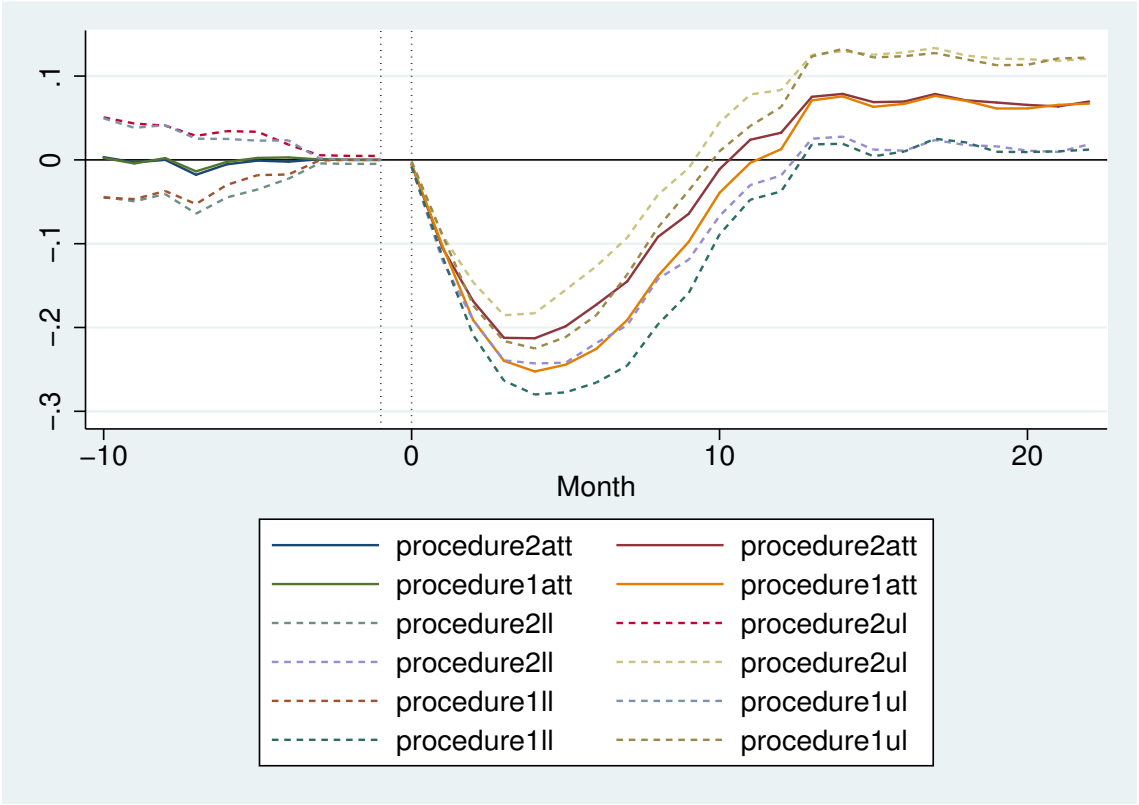
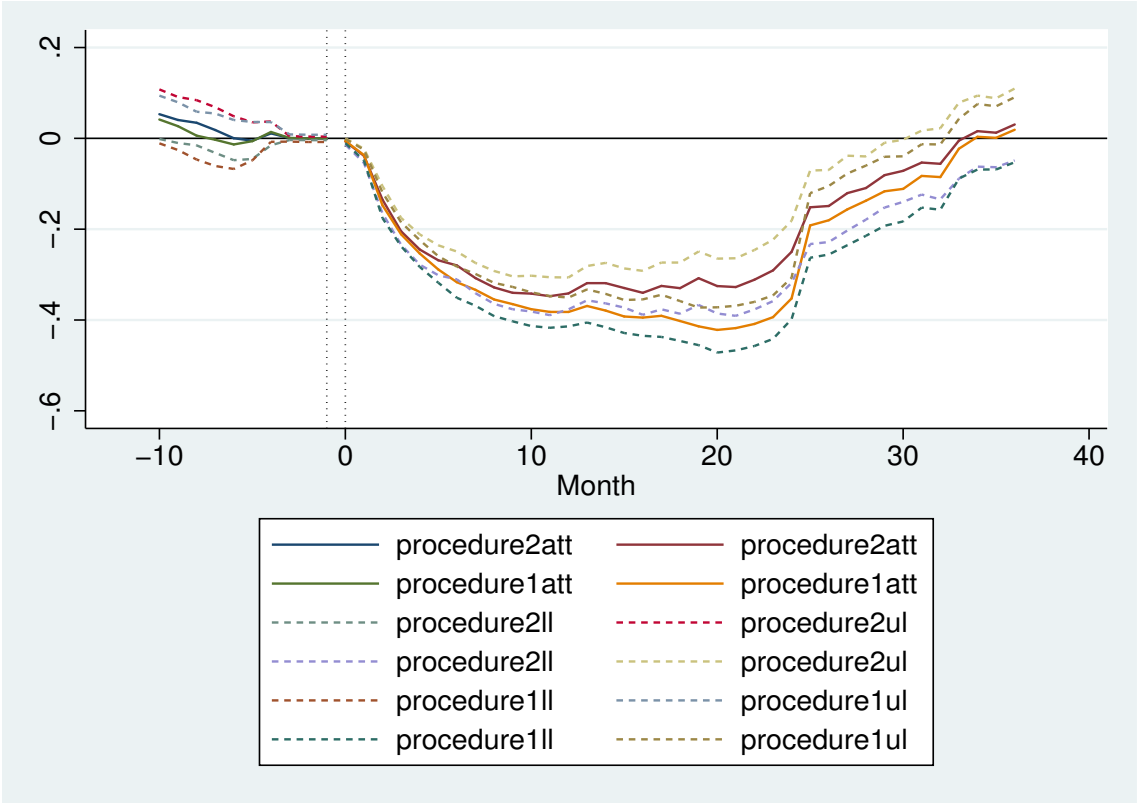
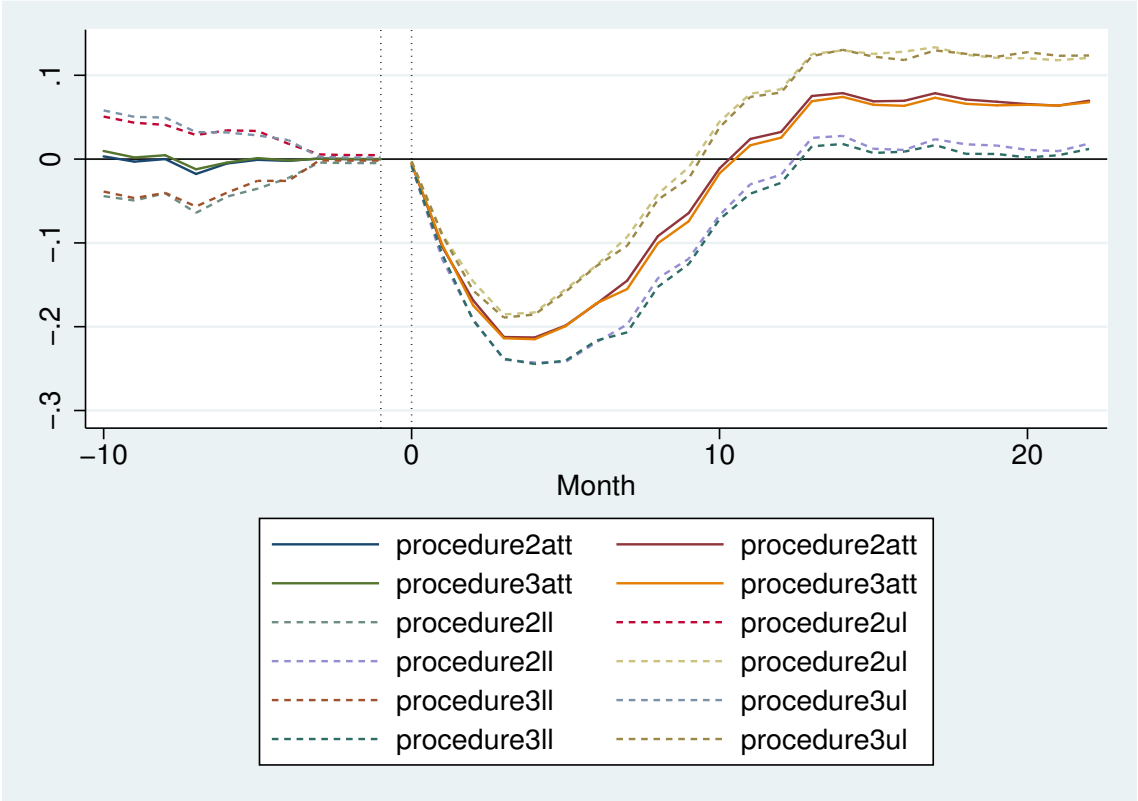


Figure 9: Treatment Effect for RT Participants with Procedure 1 and Procedure 2



(2.75 percent points when the treatment effect is positive for the first time for FT and 1.24 percent points for RT respectively). This is because the difference is mostly due to different priorities in the measurement of the outcomes and these are irrelevant once the programs have finished even according to the wrong end dates (except if participants start new programs). The bias in the treatment effect during the lock-in period caused by procedure 1, is for instance relevant if some aggregated treatment effect is calculated for cost benefit analysis.

Figure 10: Treatment Effect for FT Participants with Procedure 2 and Procedure 3



Figures 10 and 11 show how the results using procedure 3 differ from using procedure 2. The differences for FT are negligibly small (maximum 1 percentage point at a treatment effect of -14.51% in month 7). For RT they are a little larger (maximum 2.41 percentage points at a treatment effect of -31.89 % in month 13). The treatment effect is smaller using procedure 3, indicating that those who drop out because of corrections are on average more successful. This selection effect is also visible for the time before the start of the considered unemployment spell.

Procedure 3 and Procedure 4 lead to (almost) no difference for the FT results and to a small difference for RT (about 1.4 percentage points maximum). The direction of the difference is as expected, those not in the treatment group in procedure 3 compared to procedure 4 are successful per definition.

Figure 11: Treatment Effect for RT Participants with Procedure 2 and Procedure 3

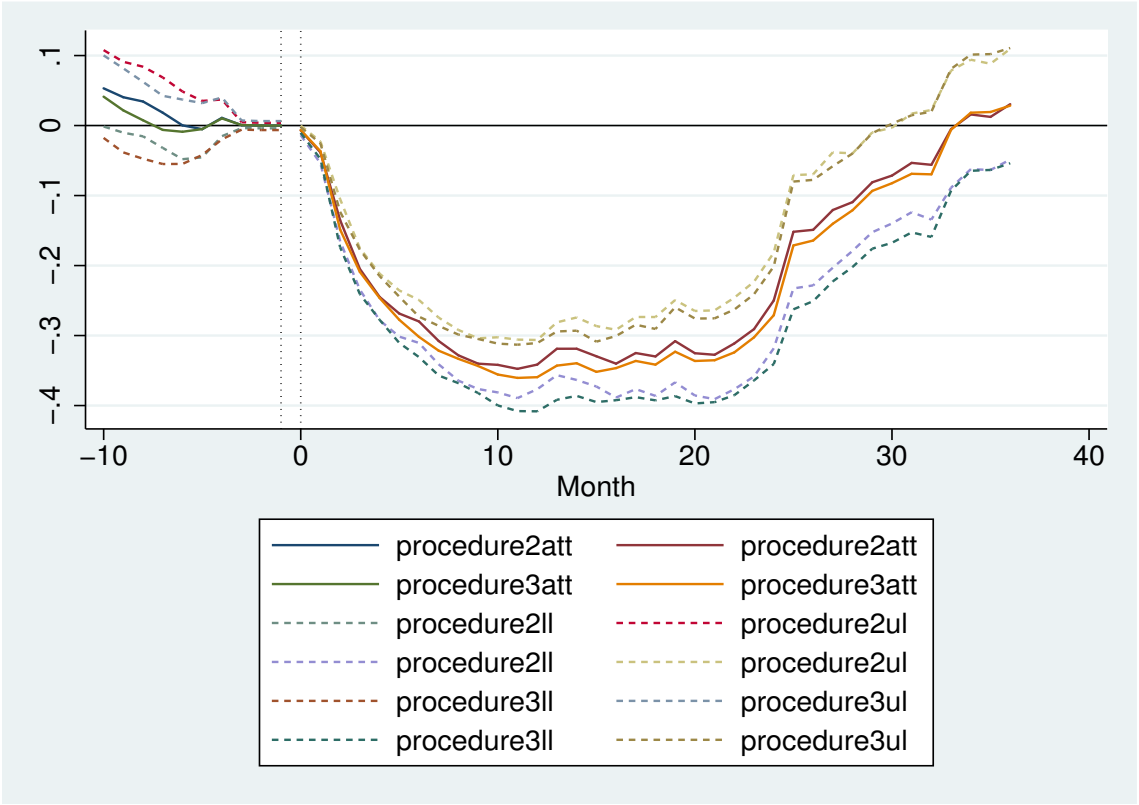


Figure 12: Treatment Effect for FT Participants with Procedure 3 and Procedure 4

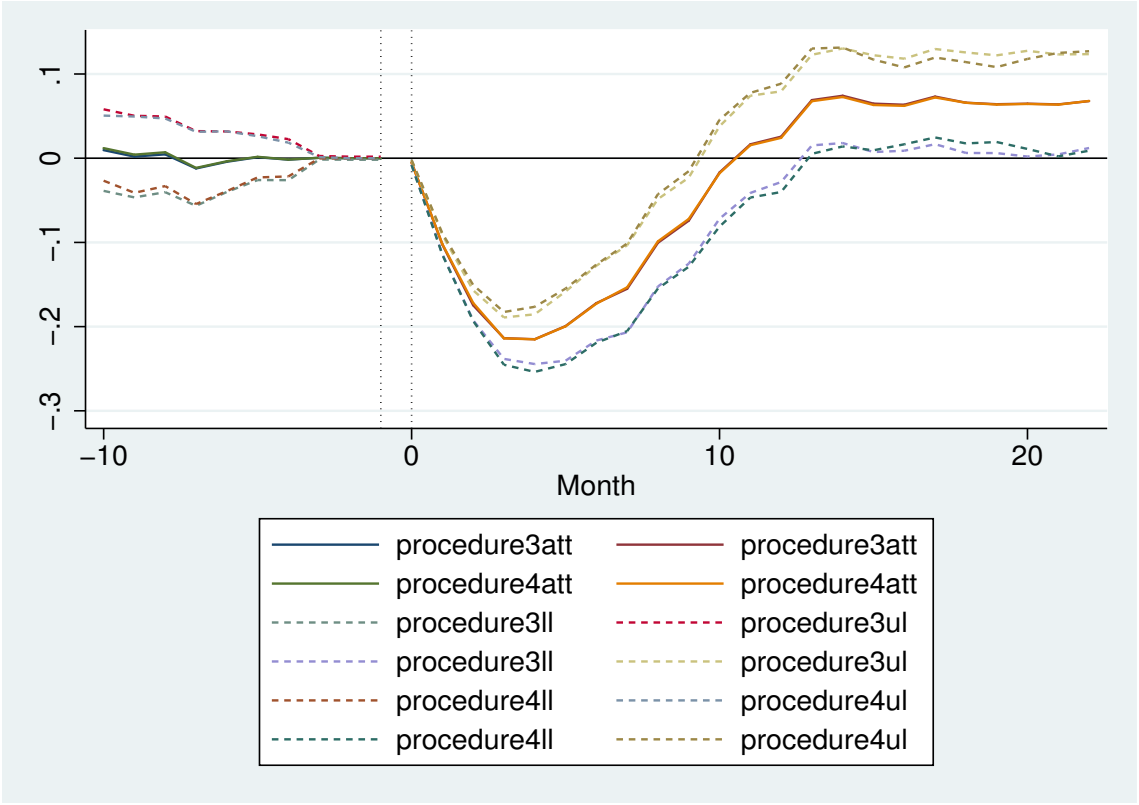
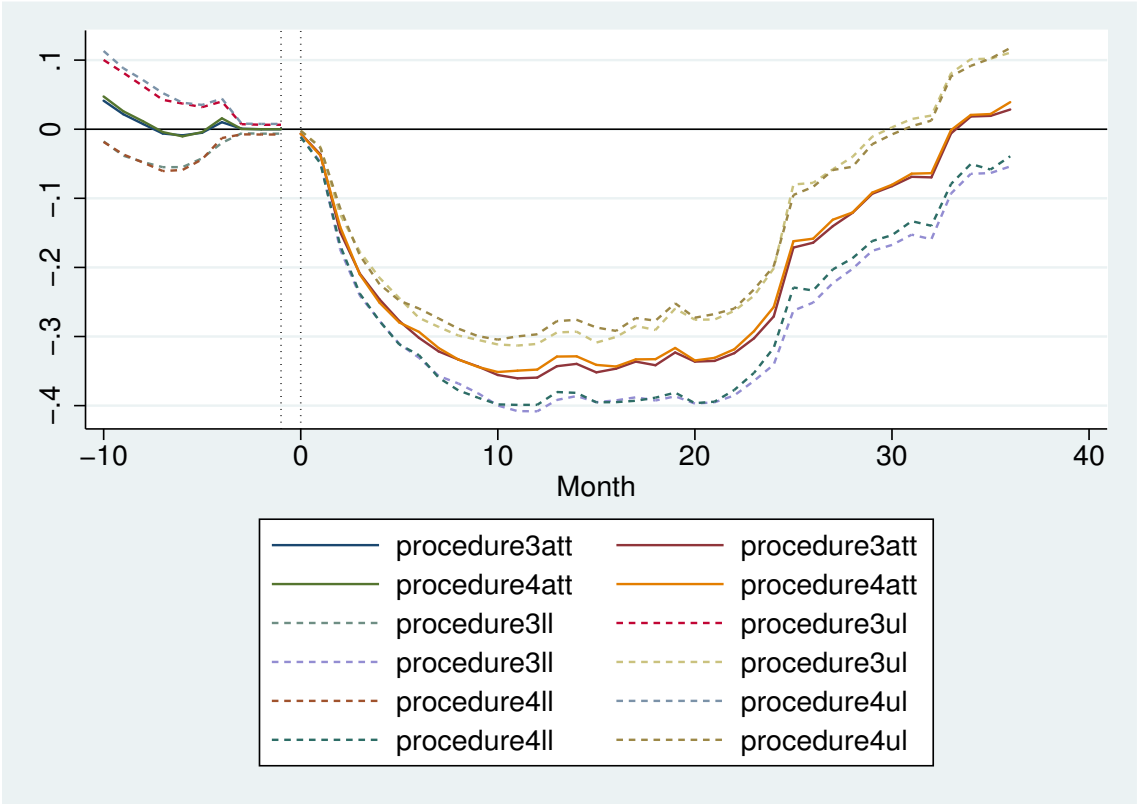


Figure 13: Treatment Effect for RT Participants with Procedure 3 and Procedure 4



As a conclusion, using propensity score matching conditional on the elapsed duration of unemployment, measurement error in end dates of programs is of relatively little importance, because it influences results only indirectly as discussed in section 4. An explicit correction of program spells might be important only for studies which need the exact magnitude of the treatment effect. But concerning the measurement of the outcome, employment spells should be given priority to program spells (procedure 1 should not be used), otherwise the magnitude of results may be considerably biased.

6 Impact Using Descriptive Duration Analysis

Apart from matching methods, duration models are very popular for the estimation of program effects. In this section a simple descriptive model is applied to analyze if the different procedures matter in a duration framework. Focussing on the duration from the start of unemployment until the start of new regular employment, a proportional hazard model with a Weibull specification is used. The sample consists of individuals who participate in the respective program during the first year of their unemployment spell and those who do not participate in any program during the

first year.[27] In addition to personal and regional characteristics and information on the individual's labor market history that are supposed to influence the hazard rate (see appendix for the final specification) three time-varying covariates are included in the estimation:[28] The day an individual enters the program under consideration the dummy variable "lock" changes to one. Once he leaves a completed program (defined as having participated at least 80% of the planned duration), the "lock" dummy changes to zero again and a second dummy ("treatfin") is set to one, indicating that this individual has finished a program.[29] In case the individual leaves an uncompleted program, "lock" is also set to zero and a third dummy ("postdrop") is set to one, indicating that the individual has dropped out of a program in the past.[30] Three time-varying dummies (instead of a simple program dummy) are included to study separately how programs bind the unemployed and the time after a completed program. The coefficients may not be interpreted as treatment effects, they just describe some aspects of the complex process that is going on. Particular problems preventing a causal interpretation are the potential endogeneity of the program end date and the relation between the dummies "lock" and "treatfin".

To investigate the impact of the correction procedure, this analysis is of interest, because the program end date is more important for the estimation than in the analysis of section 4.2 and 5, where the difference between procedure 2, 3 and 4 is only relevant for the question, if a treatment is valid. In the duration framework presented above, influence of the end date of a program on the results is still indirect, as the end date itself is neither regressor nor outcome variable. But measurement error in the end date may lead to measurement error in the covariates (for some days or for the remaining duration in case a program is wrongly assessed as having been completed), the coefficients of which shall interpreted.[31]

---

[27] For the definition of program participation see section 3.1.

[28] A time-varying covariate is interpreted as a measure of the effect of a one unit change in the covariate at time t on the log hazard (see Lancaster (1990)).

[29] The last day of a completed program is already considered as "treatfin" (if the individual leaves directly to employment), because regarding the effect of a finished program, starting a job directly after a completed program or having days of unemployment in between, is considered the same, as long as the length of the unemployment duration is in the model.

[30] This idea is inspired by Schneider et al. (2005), who distinguish between a lock-in-effect and a post program effect.

[31] A Cox model would be less suitable for the intention of this study, because time-varying covariates play only a role for the estimation at failure and not during the unemployment duration. Thus, for instance a too late switch of "lock" from one to zero would not matter, in case the individual does not leave unemployment in between.

Table 2: Extract of the results of the PH model, hazard ratios for time-varying dummies

Procedure	1	2	3	4
lock Men FT	0.061***	0.090***	0.163***	0.144***
	(1043)	(1055)	(1048)	(1051)
	(30)	(44)	(77)	(68)
lock Women FT	0.086***	0.129***	0.291***	0.261***
	(879)	(896)	(884)	(888)
	(27)	(39)	(86)	(77)
lock Men RT	0.004***	0.010***	0.036***	0.0128***
	(361)	(370)	(342)	(357)
	(3)	(6)	(18)	(7)
lock Women RT	0.006***	0.034***	0.073***	0.051***
	(313)	(314)	(299)	(308)
	(2)	(10)	(21)	(15)
treatfin Men FT	1.213***	1.241***	1.174***	1.204***
	(943)	(931)	(841)	(853)
treatfin Women FT	1.829***	1.831***	1.669***	1.718***
	(811)	(814)	(731)	(744)
treatfin Men RT	2.050***	2.371***	2.055***	2.425***
	(342)	(348)	(300)	(325)
treatfin Women RT	4.585***	5.524***	4.434***	4.987***
	(294)	(287)	(248)	(263)
postdrop Men FT	0.991	0.142	0.879	0.885
	(79)	(80)	(130)	(130)
postdrop Women FT	0.711	0.794	0.745*	0.750*
	(41)	(43)	(67)	(67)
postdrop Men RT	1.645	1.336	0.699	0.766
	(16)	(16)	(24)	(25)
postdrop Women RT	0.672	0.716	1.126	1.128
	(17)	(17)	(30)	(30)

*significant at 10% level, **significant at 5% level, ***significant at 1% level. Significance relates to being different from zero and does not mean significance of differences between the procedures. The numbers in brackets are the number of individuals that are in this state for at least one day of their duration. The second brackets of "lock" give the number of individuals who do only reach "lock", that is leave to employment (or are censored) out of an unfinished treatment. The whole number of individuals varies from about 16000 to 27000.

Table 2 shows the hazard ratios (the exponentiated coefficients) for the dummies "lock", "treatfin" and "postdrop" for men and women, FT and RT programs for West Germany for the four procedures. For the coefficients, including those of the additional covariates and standard errors see appendix. A hazard ratio of 0.09 for "lock" means, that the hazard rate for those being currently in an unfinished program is just 9% of the hazard rate of those not being in a program. As one would expect, "lock" always has a negative and highly significant effect, attending a non finished program comes along with a drastic reduction in leaving unemployment. This is also visible from the numbers in the brackets. Whereas 1048 men enter an FT program (procedure 3), only 77 end their duration out of the uncompleted program. Using the procedures with less or no corrections, much less individuals are assessed to end their duration out of an unfinished program. This influences the hazard ratios of "lock": they differ up to 4.3 percentage points between procedure 1 and 2 and up to 16.2 percentage points between procedure 2 and 3. Thus the difference between procedure 2 and 3 is more important than between procedure 2 and 1.

The large majority of those assessed to take a program finish it and "treatfin" has always a significant positive effect on the hazard rate. As discussed above, this is not to be interpreted as a positive treatment effect, it just says that individuals having finished a program leave unemployment more often than others. The hazard ratios differ a lot between the procedures, for Women RT the hazard ratio is 4.585 for procedure 1, 5.524 for procedure 2 and 4.434 for procedure 3, again the difference is larger for procedure 2 and 3 than for 1 and 2. The reason is, that a procedure without a correction mechanism misclassifies individuals to have finished a program, while they should be classified as being unemployed after an unfinished program or leaving to employment out of an unfinished program (as one can also see from the numbers in brackets). A second effect is, that in procedure 2 too many individuals are assessed as leaving directly out of an unfinished program, while in reality they have left the program even before and should be classified to "postdrop" equal to one and "lock" equal to zero. This leads c.p. to a too high hazard ratio of "lock" and a too low hazard ratio for "postdrop" using procedure 2. The coefficients of "postdrop" are not significant.

In sum, the results show that measurement error in end dates has a larger effect on results in a framework, in which the channel through which measurement error influences results, is more important than in the matching approach of section 5. In the above duration framework the end date affects the results, because it is of importance if a program has been completed and also if someone starts employment

out of an unfinished program or some time after having dropped out. Measurement error in end dates changes the magnitude but not the direction of the results.

7 Conclusion

This paper studies the error-proneness of end dates of further training programs in the IEBS with two aims: to gain insights on how to handle this problem in future studies using the IEBS and on the more general question how measurement error in end dates of treatments affects evaluation results. Mainly due to early-drop out not corrected in the data, a considerable part of end dates of further training programs are later than the actual end of participation. But the IEBS includes information that can be used to correct these end dates. In this paper four procedures how to deal with the error-prone end dates are presented, a "naive" procedure, a standard procedure and two slightly different correction mechanisms. The influence of the different procedures on evaluation results is studied using descriptive attendance and employment rates, matching and a simple descriptive duration model. This analysis shows that upward measurement error in end dates influences evaluation results through different channels, but only indirectly. Because the influence is indirect, it has only minor effects on the results if matching is used. There is almost no effect of error-prone end dates on treatment effects after the end of the program but a considerable effect on the size of the negative employment effect during the lock-in period and in particular for very long programs. This bias for the lock-in effect may for example be relevant if one is interested in averaging treatment effects for cost benefit analysis. The effect of measurement error in the end dates is larger, but does not change the direction of the results, if a duration framework with a distinction between the time in an uncompleted program and after a program, is used. The overall small effect of error prone end dates on evaluation results is good news for researchers using administrative data sets which are likely to suffer from similar problems without having the advantage to correct end dates. The advice for future users of the IEBS is to avoid using the so called "naive" procedure which gives priority to program data. An explicit correction of end dates does not seem necessary for standard evaluation, except if interest lies in the exact size of the lock-in effect.

References

Battistin, E. and B. Sianesi (2006) "Misreported schooling and returns to education: evidence from the UK" Cemmap Working Papers, CWP 07/06.

Bender, S., M. Biewen, B. Fitzenberger, M. Lechner, S. Lischke, R. Miquel, A. Osikominu, T. Wenzel, C. Wunsch (2004) "Die Beschäftigungswirkung der FbW-Maßnahmen 2000 - 2002 auf individueller Ebene: Eine Evaluation auf Basis der prozessproduzierten Daten des IAB – Vorläufiger Zwischenbericht Oktober 2004", Goethe University Frankfurt and SIAW St. Gallen.

Bender, S., M. Biewen, B. Fitzenberger, M. Lechner, R. Miquel, A. Osikominu, M. Waller, C. Wunsch (2005) "Die Beschäftigungswirkung der FbW-Maßnahmen 2000 - 2002 auf individueller Ebene: Eine Evaluation auf Basis der prozessproduzierten Daten des IAB – Zwischenbericht Oktober 2005", Goethe University Frankfurt and SIAW St. Gallen.

Biewen, M., B. Fitzenberger, A. Osikominu, M. Waller (2006) "Employment Effects of Short and Medium Term Further Training Programs in Germany in the Early 2000s", Goethe University Frankfurt

Bonin, H. and H. Schneider (2006) "Wirksamkeit der Förderung der beruflichen Weiterbildung vor und nach den Hartz-Reformen" IZA Discussion Paper No. 2069.

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

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.

Hummel, E., P. Jacobebbinghaus, A. Kohlmann, M. Oertel, C. Wübbeke, M. Ziegerer (2005) "Stichprobe der Integrierten Erwerbsbiographien, IEBS 1.0", FDZ Datenreport 6/2005, Institut für Arbeitsmarkt- und Berufsforschung (IAB), Nürnberg.

Jaenichen, U., T. Kruppe, G. Stephan, B. Ulrich and F. Wießner (2005) "You can split it if you really want: Korrekturvorschläge für ausgewählte Inkonsistenzen in der MTG und IEB", FDZ Datenreport 4/2005, Bundesagentur für Arbeit, Nürnberg.

Kruppe, M and T. Oertel, (2003) "Von Verwaltungsdaten zu Forschungsdaten", IAB Werkstattbericht 10/2003, Institut für Arbeitsmarkt- und Berufsmarktforschung (IAB), Nürnberg.

Lancaster, T. (1990) "The Econometric Analysis of Transition Data", Cambridge University Press.

Lechner, M. (2001) "Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption" In: M. Lechner and F. Pfeiffer, editors, (2000), *Econometric Evaluation of Active Labor Market Politics in Europe*, Heidelberg: Physica-Verlag.

Lechner, M. und C. Wunsch (2006) "Active Labour Market Policy in East Germany: Waiting for the Economy to Take Off." IZA Discussion Paper No. 2363.

Lewbel, A. (2004) "Estimation of Average Treatment Effect with Missclassification", Boston College.

Molinari, F. (2005) "Missing Treatments" CAE Working Paper, 05-11.

Schneider, H., K. Brenke, D. Hess, L. Kaiser, J. Steinwede, A. Uhlendorff (2005) "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 Programmes 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.

Appendix

Table 3: Variables Used in the Final PH Specification

Name	Definition
lock	whether an individual is currently attending the program under consideration
posttreat	whether a person has finished his participation in the program under consideration
agegroup1	age in between 25 and 29 years
agegroup2	age in between 30 and 34 years
agegroup3	age in between 35 and 39 years
agegroup4	age in between 40 and 44 years
countemp	number of days in regular employment within the last three years before the beginning of unemployment
lnwaged	log of daily wage in the last job(s) before the beginning of the unemployment spell, zero if above or below social security threshold
dcountub	whether receipt of unemployment benefit within the last three years before the beginning of unemployment
dcountua	whether receipt of unemployment assistance within the last three years before the beginning of unemployment
countoos	number of days out of sample within the last three years before the beginning of unemployment
countsub	number of days receiving subsistence allowance within the last three years before the beginning of unemployment
ur_qb	unemployment rate in the individual's home district in the calendar year before the beginning of unemployment
foreigner	citizenship not German
region2 to region5	classification of the districts of residence according to local labor market conditions in 5 groups (four filled for West Germany)
health2	health problems, but considered without impact on placement

Name	Definition
health3	health problems, considered to have an impact on placement
qualification1	no degree
qualification2	vocational training degree
schooling2	Hauptschulabschluss or Mittlere Reife /Fachoberschule (degrees reached after completion of the 9th or 10th grade)
schooling3	Fachhochschulreife or Abitur/Hochschulreife (degrees reached after completion of the 12th or 13th grade)
family2	living alone
family3	not married, but living together with at least one person
family4	single parent
family5	married
child	at least one child
parttime	person worked less than full-time in the last job
industry1 to industry6	industry of last employment in 6 categories
whitecollar2	previous employment was a white-collar job
occchange2	person wishes to work in the same occupation as in the last employment
quarter1 to quarter9	quarter of the calendar of the end of the last employment from 2/2000 to 1/2002
motivationlack	within the last three years there is information, that the person did not appear regularly at the labor office, on lack of cooperation, availability or similar
problemgroup	participation in a program with a social work component within the last three years

Table 4: Coefficients of PH Model Procedure 1

	Men FT	Women FT	Men RT	Women RT
lock	-2.804 (0.183)***	-2.451 (0.193)***	-5.525 (0.707)***	-5.051 (0.707)***
treatfin	0.193 (0.043)***	0.604 (0.046)***	0.723 (0.083)***	1.523 (0.085)***
postdrop	-0.009 (0.141)	-0.342 (0.230)	0.498 (0.334)	-0.397 (0.409)
agegroup1	0.517 (0.023)***	0.169 (0.031)***	0.491 (0.024)***	0.129 (0.032)***
agegroup2	0.448 (0.022)***	0.156 (0.030)***	0.433 (0.023)***	0.125 (0.031)***

	Men FT	Women FT	Men RT	Women RT
agegroup3	0.326 (0.023)***	0.198 (0.030)***	0.325 (0.023)***	0.203 (0.030)***
agegroup4	0.256 (0.025)***	0.280 (0.030)***	0.260 (0.025)***	0.268 (0.031)***
countemp	0.001 (0.000)***	0 (0.000)***	0.001 (0.000)***	0 (0.000)***
lnwaged	0.040 (0.010)***	0.051 (0.015)***	0.046 (0.010)***	0.051 (0.015)***
dcountub	0.315 (0.017)***	0.328 (0.023)***	0.323 (0.017)***	0.322 (0.024)***
dcountua	-0.178 (0.024)***	-0.152 (0.034)***	-0.166 (0.025)***	-0.118 (0.034)***
countoos	0.0003 (0.000)***	0.00009 (0.000)	0.0003 (0.000)***	0.0001 (0.000)*
countsub	0.0002 (0.000)*	0.0002 (0.000)	0.0002 (0.000)*	0.0002 (0.000)
ur_qb	-1.544 (0.187)***	-1.175 (0.262)***	-1.539 (0.189)***	-1.078 (0.264)***
foreigner	-0.108 (0.021)***	-0.109 (0.033)***	-0.105 (0.021)***	-0.114 (0.033)***
region3	0.192 (0.021)***	0.038 (0.027)	0.197 (0.021)***	0.032 (0.027)
region4	0.222 (0.027)***	0.153 (0.034)***	0.227 (0.027)***	0.142 (0.035)***
region5	0.377 (0.022)***	0.205 (0.029)***	0.387 (0.023)***	0.222 (0.030)***
health2	-0.302 (0.037)***	-0.298 (0.047)***	-0.292 (0.037)***	-0.308 (0.048)***
health3	-0.521 (0.038)***	-0.438 (0.053)***	-0.529 (0.039)***	-0.454 (0.054)***
qualification1	-0.203 (0.044)***	-0.038 (0.048)	-0.193 (0.046)***	-0.029 (0.049)
qualification2	-0.093 (0.042)**	-0.018 (0.044)	-0.068 (0.043)	0.006 (0.045)
schooling2	0.079 (0.024)***	0.036 (0.040)	0.055 (0.024)**	0.024 (0.040)
schooling3	0.026 (0.036)	0.115 (0.048)**	0.016 (0.036)	0.114 (0.049)**
family2	0.604 (0.024)***	0.843 (0.032)***	0.632 (0.024)***	0.887 (0.032)***
family3	0.616 (0.045)***	0.709 (0.055)***	0.640 (0.046)***	0.706 (0.056)***
family4	0.385 (0.096)***	0.608 (0.044)***	0.351 (0.099)***	0.624 (0.045)***
family5	0.676 (0.025)***	0.546 (0.032)***	0.691 (0.026)***	0.567 (0.032)***
child	0.192 (0.018)***	0.188 (0.025)***	0.193 (0.018)***	0.190 (0.026)***
parttime	0.050 (0.037)	-0.113 (0.024)***	0.043 (0.038)	-0.118 (0.025)***
industry1	0.229 (0.036)***	0.251 (0.073)***	0.234 (0.037)***	0.265 (0.073)***
industry2	-0.176 (0.020)***	-0.172 (0.030)***	-0.157 (0.020)***	-0.177 (0.031)***
industry4	-0.020 (0.019)	0.078 (0.024)***	-0.025 (0.019)	0.076 (0.024)***
industry5	-0.011 (0.023)	0.015 (0.028)	0.011 (0.023)	0.042 (0.029)
whitecollar2	0.182 (0.021)***	0.083 (0.026)***	0.161 (0.021)***	0.060 (0.026)**
occchange2	0.290 (0.018)***	0.208 (0.026)***	0.293 (0.019)***	0.220 (0.026)***
quarter1	-0.058 (0.031)*	-0.074 (0.043)*	-0.075 (0.031)**	-0.053 (0.044)
quarter2	-0.256 (0.030)***	-0.148 (0.038)***	-0.243 (0.030)***	-0.158 (0.039)***
quarter3	-0.278 (0.028)***	-0.129 (0.035)***	-0.266 (0.028)***	-0.119 (0.035)***
quarter4	-0.125 (0.023)***	-0.053 (0.033)	-0.117 (0.023)***	-0.048 (0.033)

	Men FT	Women FT	Men RT	Women RT
quarter5	-0.001 (0.022)	-0.073 (0.033)**	-0.007 (0.022)	-0.059 (0.033)*
quarter6	-0.258 (0.029)***	-0.116 (0.036)***	-0.249 (0.029)***	-0.122 (0.037)***
quarter7	-0.320 (0.028)***	-0.128 (0.035)***	-0.325 (0.028)***	-0.122 (0.036)***
motivationlack	-0.116 (0.023)***	-0.044 (0.032)	-0.104 (0.023)***	-0.060 (0.033)*
problemgroup	-0.046 (0.044)	-0.166 (0.068)**	-0.057 (0.044)	-0.165 (0.069)**
_cons	-6.259 (0.096)***	-5.568 (0.122)***	-6.292 (0.097)***	-5.594 (0.124)***

Table 5: Coefficients of PH Model Procedure 2

	Men FT	Women FT	Men RT	Women RT
lock	-2.398 (0.151)***	-2.048 (0.161)***	-4.582 (0.447)***	-3.389 (0.317)***
treatfin	0.216 (0.043)***	0.605 (0.046)***	0.863 (0.079)***	1.709 (0.083)***
postdrop	0.031 (0.138)	-0.231 (0.219)	0.290 (0.354)	-0.334 (0.409)
agegroup1	0.520 (0.023)***	0.170 (0.031)***	0.492 (0.024)***	0.124 (0.032)***
agegroup2	0.452 (0.022)***	0.161 (0.030)***	0.437 (0.023)***	0.125 (0.030)***
agegroup3	0.327 (0.023)***	0.203 (0.029)***	0.326 (0.023)***	0.202 (0.030)***
agegroup4	0.260 (0.024)***	0.290 (0.030)***	0.265 (0.025)***	0.276 (0.031)***
countemp	0.001 (0.000)***	0 (0.000)***	0.001 (0.000)***	0 (0.000)***
lnwaged	0.050 (0.010)***	0.050 (0.015)***	0.055 (0.010)***	0.052 (0.015)***
dcountub	0.319 (0.017)***	0.324 (0.023)***	0.327 (0.017)***	0.322 (0.023)***
dcountua	-0.183 (0.024)***	-0.146 (0.033)***	-0.169 (0.024)***	-0.120 (0.034)***
countoos	0.0003 (0.000)***	0.00009 (0.000)	0.0003 (0.000)***	0.0001 (0.000)*
countsub	0.0002 (0.000)	0.0002 (0.000)	0.0002 (0.000)	0.0002 (0.000)
ur_qb	-1.921 (0.186)***	-1.570 (0.262)***	-1.944 (0.188)***	-1.604 (0.264)***
foreigner	-0.107 (0.021)***	-0.110 (0.033)***	-0.102 (0.021)***	-0.106 (0.033)***
region3	0.183 (0.020)***	0.029 (0.027)	0.188 (0.021)***	0.023 (0.027)
region4	0.206 (0.027)***	0.139 (0.034)***	0.210 (0.027)***	0.127 (0.035)***
region5	0.359 (0.022)***	0.185 (0.029)***	0.370 (0.022)***	0.200 (0.030)***
health2	-0.290 (0.036)***	-0.291 (0.046)***	-0.281 (0.037)***	-0.311 (0.048)***
health3	-0.520 (0.038)***	-0.435 (0.052)***	-0.528 (0.038)***	-0.463 (0.053)***
qualification1	-0.206 (0.044)***	-0.041 (0.048)	-0.206 (0.045)***	-0.037 (0.049)
qualification2	-0.099 (0.042)**	-0.023 (0.043)	-0.082 (0.043)*	-0.004 (0.044)
schooling2	0.073 (0.024)***	0.038 (0.040)	0.051 (0.024)**	0.025 (0.040)
schooling3	0.025 (0.036)	0.118 (0.048)**	0.017 (0.036)	0.114 (0.049)**
family2	0.591 (0.023)***	0.833 (0.032)***	0.620 (0.024)***	0.877 (0.032)***

	Men FT	Women FT	Men RT	Women RT
family3	0.593 (0.044)***	0.694 (0.055)***	0.617 (0.045)***	0.690 (0.056)***
family4	0.355 (0.095)***	0.596 (0.044)***	0.324 (0.098)***	0.603 (0.044)***
family5	0.667 (0.025)***	0.538 (0.031)***	0.681 (0.025)***	0.552 (0.032)***
child	0.192 (0.018)***	0.184 (0.025)***	0.192 (0.018)***	0.189 (0.025)***
parttime	0.048 (0.037)	-0.106 (0.024)***	0.041 (0.037)	-0.117 (0.024)***
industry1	0.236 (0.036)***	0.253 (0.073)***	0.245 (0.036)***	0.268 (0.073)***
industry2	-0.172 (0.019)***	-0.170 (0.030)***	-0.153 (0.020)***	-0.175 (0.030)***
industry4	-0.014 (0.019)	0.080 (0.024)***	-0.018 (0.019)	0.081 (0.024)***
industry5	-0.001 (0.023)	0.015 (0.028)	0.021 (0.023)	0.039 (0.029)
whitecollar2	0.181 (0.021)***	0.088 (0.026)***	0.162 (0.021)***	0.058 (0.026)**
occchange2	0.290 (0.018)***	0.211 (0.025)***	0.292 (0.019)***	0.217 (0.026)***
quarter1	-0.047 (0.031)	-0.062 (0.043)	-0.064 (0.031)**	-0.042 (0.044)
quarter2	-0.241 (0.029)***	-0.142 (0.037)***	-0.231 (0.030)***	-0.151 (0.038)***
quarter3	-0.270 (0.028)***	-0.122 (0.035)***	-0.259 (0.028)***	-0.105 (0.035)***
quarter4	-0.120 (0.022)***	-0.057 (0.032)*	-0.114 (0.023)***	-0.055 (0.033)*
quarter5	0.002 (0.022)	-0.065 (0.032)**	-0.005 (0.022)	-0.054 (0.033)
quarter6	-0.249 (0.028)***	-0.112 (0.036)***	-0.243 (0.029)***	-0.111 (0.036)***
quarter7	-0.319 (0.027)***	-0.127 (0.035)***	-0.324 (0.028)***	-0.113 (0.035)***
motivationlack	-0.109 (0.023)***	-0.047 (0.032)	-0.097 (0.023)***	-0.063 (0.032)*
problemgroup	-0.049 (0.040)	-0.123 (0.060)**	-0.058 (0.041)	-0.085 (0.061)
_cons	-6.247 (0.096)***	-5.518 (0.121)***	-6.277 (0.097)***	-5.518 (0.124)***

Table 6: Coefficients of PH Model Procedure 3

	Men FT	Women FT	Men RT	Women RT
lock	-1.814 (0.114)***	-1.235 (0.108)***	-3.298 (0.243)***	-2.621 (0.219)***
treatfin	0.161 (0.046)***	0.512 (0.049)***	0.720 (0.088)***	1.490 (0.093)***
postdrop	-0.129 (0.112)	-0.291 (0.175)*	-0.358 (0.317)	0.119 (0.268)
agegroup1	0.517 (0.023)***	0.169 (0.031)***	0.492 (0.024)***	0.125 (0.032)***
agegroup2	0.449 (0.022)***	0.158 (0.030)***	0.436 (0.023)***	0.121 (0.030)***
agegroup3	0.325 (0.023)***	0.200 (0.029)***	0.325 (0.023)***	0.200 (0.030)***
agegroup4	0.256 (0.024)***	0.285 (0.030)***	0.261 (0.025)***	0.272 (0.031)***
countemp	0.001 (0.000)***	0 (0.000)***	0.001 (0.000)***	0 (0.000)***
lnwaged	0.049 (0.010)***	0.050 (0.015)***	0.054 (0.010)***	0.052 (0.015)***
dcountub	0.319 (0.017)***	0.322 (0.023)***	0.326 (0.017)***	0.321 (0.023)***

	Men FT	Women FT	Men RT	Women RT
dcountua	-0.179 (0.024)***	-0.140 (0.033)***	-0.164 (0.024)***	-0.123 (0.034)***
countoos	0.0003 (0.000)***	0.00008 (0.000)	0.0003 (0.000)***	0.0001 (0.000)
countsub	0.0002 (0.000)*	0.0002 (0.000)	0.0002 (0.000)*	0.0002 (0.000)
ur_qb	-1.507 (0.185)***	-1.752 (0.262)***	-1.585 (0.187)***	-1.689 (0.264)***
foreigner	-0.108 (0.021)***	-0.108 (0.033)***	-0.105 (0.021)***	-0.103 (0.033)***
region3	0.188 (0.020)***	0.029 (0.027)	0.192 (0.021)***	0.025 (0.027)
region4	0.215 (0.027)***	0.134 (0.034)***	0.216 (0.027)***	0.124 (0.035)***
region5	0.374 (0.022)***	0.184 (0.029)***	0.382 (0.022)***	0.197 (0.030)***
health2	-0.288 (0.036)***	-0.294 (0.046)***	-0.280 (0.037)***	-0.314 (0.048)***
health3	-0.515 (0.038)***	-0.439 (0.052)***	-0.524 (0.038)***	-0.465 (0.053)***
qualification1	-0.205 (0.044)***	-0.041 (0.048)	-0.202 (0.045)***	-0.038 (0.049)
qualification2	-0.096 (0.042)**	-0.022 (0.043)	-0.079 (0.043)*	-0.005 (0.044)
schooling2	0.072 (0.024)***	0.041 (0.040)	0.053 (0.024)**	0.031 (0.040)
schooling3	0.021 (0.036)	0.120 (0.048)**	0.017 (0.036)	0.118 (0.049)**
family2	0.595 (0.023)***	0.831 (0.032)***	0.624 (0.024)***	0.876 (0.032)***
family3	0.598 (0.044)***	0.691 (0.055)***	0.618 (0.045)***	0.692 (0.056)***
family4	0.350 (0.095)***	0.591 (0.044)***	0.322 (0.098)***	0.604 (0.044)***
family5	0.668 (0.025)***	0.534 (0.031)***	0.685 (0.025)***	0.548 (0.032)***
child	0.199 (0.018)***	0.186 (0.025)***	0.197 (0.018)***	0.191 (0.025)***
parttime	0.047 (0.037)	-0.107 (0.024)***	0.040 (0.037)	-0.116 (0.025)***
industry1	0.236 (0.036)***	0.245 (0.073)***	0.244 (0.036)***	0.263 (0.073)***
industry2	-0.169 (0.019)***	-0.171 (0.030)***	-0.150 (0.020)***	-0.173 (0.030)***
industry4	-0.013 (0.019)	0.077 (0.024)***	-0.016 (0.019)	0.081 (0.024)***
industry5	-0.007 (0.023)	0.015 (0.028)	0.015 (0.023)	0.040 (0.029)
whitecollar2	0.180 (0.021)***	0.090 (0.026)***	0.163 (0.021)***	0.062 (0.026)**
occchange2	0.290 (0.018)***	0.212 (0.025)***	0.292 (0.019)***	0.218 (0.026)***
quarter1	-0.054 (0.031)*	-0.056 (0.043)	-0.069 (0.031)**	-0.038 (0.044)
quarter2	-0.244 (0.029)***	-0.141 (0.037)***	-0.232 (0.030)***	-0.146 (0.038)***
quarter3	-0.268 (0.028)***	-0.119 (0.035)***	-0.261 (0.028)***	-0.100 (0.035)***
quarter4	-0.117 (0.022)***	-0.053 (0.032)	-0.111 (0.023)***	-0.050 (0.033)
quarter5	-0.001 (0.022)	-0.060 (0.032)*	-0.005 (0.022)	-0.051 (0.033)
quarter6	-0.257 (0.028)***	-0.108 (0.036)***	-0.251 (0.029)***	-0.105 (0.036)***
quarter7	-0.322 (0.027)***	-0.126 (0.035)***	-0.326 (0.028)***	-0.114 (0.035)***
motivationlack	-0.110 (0.023)***	-0.049 (0.032)	-0.100 (0.023)***	-0.065 (0.032)**
problemgroup	-0.055 (0.040)	-0.124 (0.060)**	-0.062 (0.041)	-0.084 (0.061)

	Men FT	Women FT	Men RT	Women RT
_cons	-6.299 (0.096)***	-5.510 (0.122)***	-6.329 (0.097)***	-5.517 (0.124)***

Table 7: Coefficients of PH Model Procedure 4

	Men FT	Women FT	Men RT	Women RT
lock	-1.939 (0.122)***	-1.344 (0.115)***	-4.361 (0.408)***	-2.970 (0.259)***
treatfin	0.185 (0.045)***	0.541 (0.048)***	0.886 (0.082)***	1.607 (0.088)***
postdrop	-0.122 (0.112)	-0.288 (0.175)*	-0.267 (0.302)	0.121 (0.268)
agegroup1	0.517 (0.023)***	0.166 (0.031)***	0.492 (0.024)***	0.121 (0.032)***
agegroup2	0.449 (0.022)***	0.159 (0.030)***	0.436 (0.023)***	0.123 (0.030)***
agegroup3	0.325 (0.023)***	0.202 (0.029)***	0.325 (0.023)***	0.201 (0.030)***
agegroup4	0.257 (0.024)***	0.285 (0.030)***	0.263 (0.025)***	0.271 (0.031)***
countemp	0.001 (0.000)***	0 (0.000)***	0.001 (0.000)***	0 (0.000)***
lnwaged	0.050 (0.010)***	0.052 (0.015)***	0.055 (0.010)***	0.053 (0.015)***
dcountub	0.319 (0.017)***	0.321 (0.023)***	0.326 (0.017)***	0.319 (0.023)***
dcountua	-0.179 (0.024)***	-0.139 (0.033)***	-0.164 (0.025)***	-0.120 (0.034)***
countoos	0.0003 (0.000)***	0.00006 (0.000)	0.0003 (0.000)***	0.0001 (0.000)*
countsub	0.0002 (0.000)*	0.0001 (0.000)	0.0001 (0.000)*	0.0002 (0.000)
ur_qb	-1.535 (0.185)***	-1.629 (0.263)***	-1.575 (0.187)***	-1.593 (0.266)***
foreigner	-0.106 (0.021)***	-0.109 (0.033)***	-0.102 (0.021)***	-0.102 (0.033)***
region3	0.189 (0.020)***	0.027 (0.027)	0.194 (0.021)***	0.024 (0.027)
region4	0.216 (0.027)***	0.137 (0.034)***	0.216 (0.027)***	0.128 (0.035)***
region5	0.374 (0.022)***	0.185 (0.029)***	0.384 (0.022)***	0.199 (0.030)***
health2	-0.290 (0.036)***	-0.292 (0.046)***	-0.280 (0.037)***	-0.312 (0.048)***
health3	-0.522 (0.038)***	-0.436 (0.052)***	-0.532 (0.038)***	-0.462 (0.053)***
qualification1	-0.206 (0.044)***	-0.033 (0.048)	-0.202 (0.045)***	-0.028 (0.049)
qualification2	-0.096 (0.042)**	-0.014 (0.043)	-0.078 (0.043)*	0.006 (0.044)
schooling2	0.073 (0.024)***	0.040 (0.040)	0.053 (0.024)**	0.029 (0.040)
schooling3	0.020 (0.036)	0.121 (0.048)**	0.016 (0.036)	0.120 (0.049)**
family2	0.597 (0.023)***	0.838 (0.032)***	0.626 (0.024)***	0.883 (0.032)***
family3	0.598 (0.044)***	0.697 (0.055)***	0.617 (0.045)***	0.697 (0.055)***
family4	0.361 (0.095)***	0.597 (0.044)***	0.331 (0.098)***	0.608 (0.044)***
family5	0.672 (0.025)***	0.541 (0.031)***	0.689 (0.025)***	0.555 (0.032)***
child	0.195 (0.018)***	0.186 (0.025)***	0.192 (0.018)***	0.191 (0.025)***
parttime	0.051 (0.037)	-0.104 (0.024)***	0.044 (0.037)	-0.115 (0.024)***

	Men FT	Women FT	Men RT	Women RT
industry1	0.233 (0.036)***	0.245 (0.073)***	0.240 (0.036)***	0.263 (0.073)***
industry2	-0.172 (0.019)***	-0.171 (0.030)***	-0.154 (0.020)***	-0.174 (0.030)***
industry4	-0.016 (0.019)	0.077 (0.024)***	-0.020 (0.019)	0.080 (0.024)***
industry5	-0.007 (0.023)	0.017 (0.028)	0.015 (0.023)	0.041 (0.029)
whitecollar2	0.182 (0.021)***	0.092 (0.026)***	0.164 (0.021)***	0.062 (0.026)**
occchange2	0.288 (0.018)***	0.214 (0.025)***	0.289 (0.019)***	0.219 (0.026)***
quarter1	-0.050 (0.031)	-0.061 (0.043)	-0.067 (0.031)**	-0.046 (0.044)
quarter2	-0.246 (0.029)***	-0.146 (0.037)***	-0.234 (0.030)***	-0.153 (0.038)***
quarter3	-0.264 (0.028)***	-0.123 (0.035)***	-0.259 (0.028)***	-0.102 (0.035)***
quarter4	-0.119 (0.022)***	-0.055 (0.032)*	-0.113 (0.023)***	-0.053 (0.033)
quarter5	-0.002 (0.022)	-0.060 (0.032)*	-0.007 (0.022)	-0.049 (0.033)
quarter6	-0.255 (0.028)***	-0.110 (0.036)***	-0.249 (0.029)***	-0.108 (0.036)***
quarter7	-0.318 (0.027)***	-0.125 (0.034)***	-0.322 (0.028)***	-0.111 (0.035)***
motivationlack	-0.108 (0.023)***	-0.050 (0.032)	-0.097 (0.023)***	-0.065 (0.032)**
problemgroup	-0.049 (0.040)	-0.126 (0.060)**	-0.056 (0.041)	-0.086 (0.061)
_cons	-6.307 (0.096)***	-5.541 (0.121)***	-6.336 (0.097)***	-5.542 (0.123)***