I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10531

Comparing Econometric Methods to Empirically Evaluate Job-Search Assistance

Paul Muller
Bas van der Klaauw
Arjan Heyma

JANUARY 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10531

Comparing Econometric Methods to Empirically Evaluate Job-Search Assistance

Paul Muller
University of Gothenburg

Arjan Heyma
SEO Economic Research

Bas van der Klaauw
VU University Amsterdam, Tinbergen Institute and IZA

JANUARY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Comparing Econometric Methods to Empirically Evaluate Job-Search Assistance*

We test whether different empirical methods give different results when evaluating job-search assistance programs. Budgetary problems at the Dutch unemployment insurance (UI) administration in March 2010, caused a sharp drop in the availability of these programs. Using administrative data provided by the UI administration, we evaluate the effect of the program using (1) the policy discontinuity as a quasi-experiment, (2) conventional matching methods, and (3) the timing-of-events model. All three methods use the same data to consider the same program in the same setting, and also yield similar results. The program reduces job finding during the first six months after enrollment. At longer durations, the quasi-experimental estimates are not significantly different from zero, while the non-experimental methods show a small negative effect.

JEL Classification: J64, C14, C31

Keywords: empirical policy evaluation, job-search assistance, unemployment duration

Corresponding author:
Paul Muller
Department of Economics
University of Gothenburg
P.O.Box 640
SE 405 30 Gothenburg
Sweden

E-mail: paul.muller@gu.se

* We thank UWV, and in particular Han van der Heul for making their data available and providing information on institutions. This paper has benefited from comments and suggestions by Grégory Jolivet, Jaap Abbring, Mike Elsby, Bart Cockx, Bastian Ravesteijn and seminar participants at EALE (2013), ESPE (2014), VU University Amsterdam (2014), University of Edinburgh (2014), BeNA Berlin (2014), University of Vienna (2015), University of Copenhagen (2015), University of Gothenburg (2015), DIW Berlin (2015) IZA Bonn (2015) and Collegio Carlo Alberto Turin (2015).

1 Introduction

In 2002, the Dutch market for job-search assistance programs was privatized, implying that the unemployment insurance (UI) administration buys services of private companies to assist benefits recipients in their job search. Due to the economic crisis the demand for programs increased sharply in 2009 and early 2010, leading to budgetary problems in March 2010. The government refused to extend the budget and as a result, the purchase of new programs was terminated within a period of two weeks. During the remainder of the year, new UI benefits recipients could no longer enroll in these programs. In this paper, we exploit this policy discontinuity to evaluate the effects on job finding. In addition, we estimate the same effects using non-experimental methods (matching and timing-of-events) and compare the results of the three methods.

The main challenge in evaluating active labor-market programs is selective participation (Heckman et al. (1999), Abbring and Heckman (2007)). As shown in a meta-analysis by Card et al. (2010), over 50% of evaluation studies use longitudinal data and compare a treatment group with a control group, where the control group is typically constructed by matching on observed characteristics. Over one-third of the studies use duration models. Less than 10% of the studies exploit an experimental design. In his seminal study, LaLonde (1986) shows that non-experimental estimators produce results that do not concur with those from experimental evidence. Dehejia and Wahba (1999) test the performance of matching estimators, finding that those are closer to the experimental evidence. Their findings are, however, disputed by Smith and Todd (2005), who evaluate the same program and show that the findings are not robust to different specifications and the use of different samples and different sets of covariates. They refer to Heckman et al. (1997) who argue that matching estimators can only replicate experimental findings if three requirements are fulfilled. First, use the same data source for treatment and control group (in particular the outcome variable should be measured in the same way).

Second, treated and control individuals are active in the same local labor market. Third, the data should contain a rich set of variables that affect both program participation and labor-market outcomes. Smith and Todd (2005) argue that each of these requirements is likely to be violated in the evaluations by LaLonde (1986) and Dehejia and Wahba (1999).

We build on this literature by performing a similar comparison of methods, using administrative data. Our main contribution is twofold. First, since our quasi-experimental estimates are identified from a large-scale policy discontinuity in 2010, the setting is particularly suitable for such a comparison. Our data fulfill the criteria mentioned by Smith and Todd (2005). The administrative data allow the use of high-quality information on a rich set of variables, including individual characteristics, pre-unemployment labor-market variables, current unemployment spell characteristics and any assistance provided by the UI administration and private providers. As the policy discontinuity was nationwide, the sample size is substantial. Since the policy discontinuity occurred recently, programs and labor-market conditions are similar to those currently in many countries. Second, for the non-experimental analysis we not only apply matching estimators, but also estimate the timing-of-events model.

We exploit the policy discontinuity to estimate how program participation affects job finding. The variation in program provision due to the quasi-experiment is large. Within a month, the weekly number of new program participants dropped from 1300 to less than 80 and remained below 50 for the remainder of the year. We estimate the treatment effect on the treated by comparing the job-finding rates of cohorts entering unemployment at different points in time, though relatively short after each other. Since they reach the discontinuity at different unemployment durations they are affected differentially, which identifies the effect of the programs. Seasonal differences in the labor market are controlled for using cohorts from the previous year. Our results show that after starting a program, the job-finding rate is reduced

significantly for several months (i.e. lock-in effect). After half a year it increases, up to a zero difference in job finding 12 to 18 months after the enrollment in the program.

We compare these results to matching estimators (using inverse probability weighting). The results show a significant negative effect of program participation directly after entering the program. Even though the negative effect decreases in magnitude over time, estimated effects remain significantly negative after 18 months. Next, we estimate a timing-of-events model (Abbring and Van den Berg (2003)), which allows for selection on unobservables by adding more structure to the model. It jointly models the hazard rate to employment and the hazard rate to program participation. Estimating the model we find that the program reduces the job-finding rate in the first six months, while it slightly increases the job-finding rate at longer durations. Overall this leads to a negative effect on employment in the first two years, and a zero effect afterwards.

To summarize, all methods find a significantly negative effect in the short run, while only the quasi-experimental estimates and the timing-of-events model find that the effect in the medium run is zero. The difference in results is small, and the main policy recommendations are the same for all methods. The different methods use different samples. While matching and timing-of-events allow using a large sample containing all individuals entering UI over more than two years, the quasi-experimental approach requires focusing on a smaller sample that is most affected by the discontinuity. We test whether the matching and timing-of-events results depend on the choice of sample and find that results are almost identical when applying these methods to the smaller sample.

Interpreting the findings should be done with care. The quasi-experimental results rely on a common-trend assumption, stating that in the absence of the policy discontinuity, different cohorts experience similar outflow to work, up to a constant difference in the hazard rates. Furthermore, the job-search assistance program con-

tains a mix of caseworker meetings, job referrals and goal setting in the job-search process. These elements are often also present in job-search assistance programs in other countries, but likely with a different intensity. The job-search assistance is offered in addition to of the "basic" assistance of the UI administration (mostly irregular meetings with caseworkers), which is also the case in other countries. The set-up in which caseworkers have substantial discretion when deciding which job seekers are assigned to programs is a feature common to many UI administrations. Finally, the UI benefits system is quite generous when being compared to the US or UK, but similar to other continental European countries.

The remainder of the paper is structured as follows. We briefly discuss the literature on the evaluation of active labor-market programs in Section 2, and describe the institutional setting and the budgetary problems which led to the policy discontinuity in Section 3. An overview of the data is provided in Section 4. In Section 5 we define our treatment effect of interest. In Section 6 we present non-experimental results from the matching and timing-of-events estimators. In section 7 we discuss how the discontinuity allows identifying the same treatment effect and present estimation results. Section 8 compares the results from the different methods and provides a discussion. Section 9 concludes.

2 Literature

There is a large literature studying the effectiveness of active labor-market programs. Reviews provided by Kluve (2010) and Card et al. (2010) show that there is some consensus on the effects of different types of programs on post-unemployment outcomes such as employment, wages and job stability. For example, job-search assistance is found to have more positive effects on job finding than alternatives such as public-sector employment programs.

Card et al. (2010) report that 9% of the evaluation studies in their meta-analysis use an experimental approach.[1] While randomized experiments are often considered the gold standard, they have also been criticized, mostly on practical grounds. For example, as discussed by Heckman et al. (1999), experiments are typically expensive, difficult to implement and may lead to ethical objections. Alternatively, quasi-experimental approaches are used.[2] Such approaches use institutional features or policy changes that generate random variation in program participation. When convincing sources of variation are found, they may offer useful alternatives to randomized experiments. Such variation may only apply to specific groups, leading to complications such as small samples or local treatment effects. Many studies use non-experimental methods such as matching (over 50% of the studies listed by Card et al. (2010).[3] Especially with increasingly availability of high quality administrative data, matching approaches become attractive. Some other studies apply the timing-of-events model (Abbring and Van den Berg (2003)), which makes some functional-form assumptions which allow to model unobserved factors that lead to selective program participation.[4]

An interesting question is whether a relationship exists between the methodology and the empirical results. This question is investigated across studies, in the surveys by Card et al. (2010) and Kluve (2010). They find little evidence suggesting such

[1] See for example, Van den Berg and Van der Klaauw (2006) who analyze a randomized experiment on counseling and monitoring, to show that the program merely shifts job-search effort from the informal to the formal search channel. Graversen and Van Ours (2008) evaluate an intensive activation program in Denmark, using a randomized experiment. Card et al. (2011) show estimates of the effect of a training program offered to a random sample of applicants in the Dominican Republic. Behaghel et al. (2014) perform a large controlled experiment, randomizing job seekers across publicly and privately provided counseling programs.

[2] Dolton and O'Neill (2002) use random delays in program participation to assess the effect of a job search assistance program in the UK. Van der Klaauw and Van Ours (2013) analyze the effect of both an re-employment bonus and sanctions, exploiting policies changes in the bonus levels. Cockx and Dejemeppe (2012) use a regression-discontinuity approach to estimate the effect of extra job-search monitoring in Belgium. Van den Berg et al. (2014) apply regression discontinuity with duration data to the introduction of the New Deal for the Young People in the UK.

[3] For example, Brodaty et al. (2002) apply a matching estimator to estimate the effect of activation programs for long-term unemployed workers in France, Sianesi (2004) investigates different effects of active labor-market programs in Sweden and Lechner et al. (2011) looks at long-run effects of training programs in Germany.

[4] For example, this model is used to evaluate the effect of benefit sanctions in the Netherlands (Van den Berg et al. (2004) and Abbring et al. (2005)) and to evaluate a training program in France by Crépon et al. (2012).

a relationship. In particular, Card et al. (2010) find no significant difference in results between experimental and non-experimental studies. Ideally, one would like to compare different approaches in the same the setting.

Several studies have focused on comparing outcomes of different methods. LaLonde (1986), Dehejia and Wahba (1999) and Smith and Todd (2005) show that the difference between experimental and non-experimental estimates can be substantial, using data from a job-training program in the US. These results were followed by a number of papers that consider comparing different methodologies. Lalive et al. (2008) evaluate the effect of activation programs in Switzerland. They find that matching estimators and estimates from a timing-of-events model lead to different results. Mueser et al. (2007) use a wide set of matching estimators to estimate the earnings effect of job-training programs in the US. They compare findings of the different estimators to estimates from experimental methods reported in the literature. Biewen et al. (2014) compare estimates of the effect of training programs, showing that results are sensitive to data features and methodological choices. Kastoryano and Van der Klaauw (2011) evaluate job-search assistance and compare different methods for dynamic treatment evaluation and find that results are similar.

Despite the increased focus on comparing different methods, evidence applying both (quasi-)experimental and non-experimental methods within the same setting is scarce. As argued by Heckman et al. (1997), a valid comparison between matching estimates and experimental estimates requires applying those methods using the same program, the same geographical region and the same data source. We contribute to the literature by exploiting an administrative data set that allows for different methods while fulfilling these criteria.

3 Institutional setting and the policy discontinuity

In this section we briefly describe the institutional setting at the moment of our observation period and the policy discontinuity in program provision, which we exploit in the empirical analysis.

In the Netherlands, UI is organized at the nationwide level. The UI administration (UWV) pays benefits to workers, who involuntary lost at least five working hours per week (or half of their working hours if this is less than five). Workers should have worked for at least 26 weeks out of the last 36 weeks. Fulfillment of this "weeks condition" provides an eligibility to benefits for three months. If the worker has worked at least four out of the last five years, the benefits eligibility period is extended with one month for each additional year of employment. The maximum UI benefits duration is 38 months. During the first two months benefits are 75% of the previous wage, capped at a daily maximum. From the third month onward it is 70% of the previous wage (see De Groot and Van der Klaauw (2014) for a more extensive discussion).

A UI benefits recipient is required to register at the public employment office, and to search for work actively. The latter requires making at least one job application each week. caseworkers at the UI administration provide basic job-search assistance through individual meetings. Benefit recipients are obliged to accept any suitable job offer.[5] Caseworkers are responsible for monitoring these obligations. In general, the intensity of meetings is low though (only in case the caseworker suspects that a recipient is unable to find work without assistance a meeting is scheduled). In 2009, caseworkers had the possibility of assigning an individual to a range of programs aiming at increasing the job-finding rate, if she judged that the benefits recipient required more than the usual guidance. A large diversity of programs

[5]During the first six months a suitable job is defined as a job at the same level as the previous job, between six and 12 months it can be a job below this level, and after 12 months any job is suitable and should be accepted.

existed, including job-search assistance, vacancy referral, training in writing application letters and CV's, wage subsidies, subsidized employment in the public sector and schooling. Some of these were provided internally by the UI administration, while others were purchased externally from private companies. Our analysis focuses on the externally provided programs. These can be broadly classified as (with relative frequency in parentheses) job search assistance programs (56%), training or schooling (31%), subsidized employment (2%) and other programs (11%). Though some guidelines existed, caseworkers had a large degree of discretion in deciding about program assignment.

The lack of centralized program assignment together with an increased inflow in unemployment due to the recession caused that many more individuals were assigned to these programs in 2009 and early 2010 than the budget allowed. Therefore, the entire budget had been exhausted by March 2010. Authorities refused to extend the budget and declared that no new programs should be purchased from that moment onward.[6] Assistance offered internally by the UI administration continued without change. In Section 4 we show that indeed the number of new program entrants dropped to almost zero in March 2010 and remained very low afterwards.

4 Data

We use a large administrative dataset provided by the UI administration, containing all individuals who started collecting UI benefits between April 2008 and September 2010 in the Netherlands. The dataset contains 608,998 observations (each UI spell is considered an observation, though for some individuals there are multiple spells).[7]

[6]This was declared by the minister of social affairs, in a letter to parliament on March 15. An exception was made for a small number of specially targeted programs (mostly for long-term unemployed workers).

[7]The original dataset contains 671,743 unemployment spells. We exclude 35,671 spells from individuals previously employed in the public sector and 17,577 from individuals older than 60 years. Next, we drop 533 spells from individuals working more than 60 hours or less than 12 hours in their previous job and 151 spells from individuals who were eligible for a so-called 'education and development fund'. Finally we exclude 8518 spells with a duration of zero days and 290 spells from individuals with inconsistent or missing data (such as a negative unemployment duration). These are often individuals for whom the application to benefits was later denied.

We select a sample of individuals with a high propensity to participate in the external programs. These are native males, aged 40 to 60 years, with a low unemployment history (at most two unemployment spells in the past 3 years) and belonging to the upper 60% of the income distribution in our data. This reduces the sample to 116,866 observations. The advantage of restricting the sample is twofold. First, the policy discontinuity affects this group strongest, which strengthens the first stage of the analysis. Second, the estimates are more precise when using a homogeneous sample.[8]

For each spell we observe the day of starting receiving UI benefits and, if the spell is not right censored, the last day and the reason for the end of the benefit payments. Right censoring occurs on January 1st, 2012, when our data was constructed, so for each individual we can observe at least 16 months of benefits receipt. The dataset contains a detailed description of all activation programs (both internally and externally provided) in which benefits recipients participated. Furthermore individual characteristics and pre-unemployment labor-market outcomes are included in the dataset.

Figure 1 shows how the monthly number of individuals entering UI evolves over time. Due to the economic crisis, there is a substantial increase in the inflow from December 2008 onward. The inflow increased from about 2000 to 5000 per month and remained high until the end of 2009. From 2010 onward the inflow decreased somewhat. Table 1 presents summary statistics for the full sample, as well as for three subgroups defined by their month of inflow into unemployment. Column (1) shows that for the full sample the median duration of unemployment is 245 days (around eight months). Almost 60% of those exiting UI find work, while 15% reach the end of their benefits entitlement period. Almost 7% leave unemployment due to sickness or disability, the rest leave for other reasons or the reason for exit is unknown. Exits due to reaching the end of the entitlement period and exits due

[8]For example, using a homogeneous sample improves the performance of the matching estimators applied in Subsection 6.1 (Imbens (2014)).

Table 1: Descriptive statistics

		Inflow cohort UI:		
	Full sample	April 2008	April 2009	April 2010
Unemployment duration (median, days)	245	175	280	275
Reason for exit (%):				
Work	70.1	64.5	65.3	74.8
End of entitlement period	15.4	21.5	19.1	8.3
Sickness/Disability	6.7	6.2	7.3	7.6
Other	7.8	7.8	8.3	9.3
Participation external program (%):				
Any program	18.7	24.0	32.3	0.7
Job-search assistance	11.0	17.1	21.3	0.3
Training	6.0	7.3	9.7	0.2
Subsidized employment	0.4	0.3	0.8	0.0
Other	4.9	4.4	7.7	0.2
Participation internal program (%):				
Any program	36.8	13.5[a]	40.8	39.2
Job-search assistance	11.6	1.7[a]	12.1	13.6
Subsidized employment	3.2	1.1[a]	3.9	3.8
Tests	9.7	1.9[a]	10.3	10.9
Workshop entrepreneurship	4.4	2.3[a]	6.3	3.4
Other	19.7	9.1[a]	23.4	19.6
Gender (% males) [b]	100	100	100	100
Immigrant (%)[b]	0.0	0.0	0.0	0.0
Previous hourly wage (%):				
Low[b]	0.0	0.0	0.0	0.0
Middle	57.4	53.8	58.3	53.1
High	42.6	46.2	41.7	46.9
Age	48.7	49.0	48.6	48.9
Unemployment size (hours)[c]	37.2	37.1	37.4	37.3
UI history last 3 year (%)	29.2	33.8	28.8	23.3
Education (%):				
Low	22.8	20.0	21.6	20.5
Middle	46.5	43.0	45.7	47.1
High	30.7	37.1	32.7	32.4
Observations	116,866	1774	4441	4505

Job-search assistance contains 'IRO' (Individual reintegration agreement), 'Job hunting' and 'Application letter'. Training contains 'Short Training' and 'Schooling'. Subsidized employment contains 'Learn-work positions'. [a] Biased downwards, because participation in internal programs was rarely recorded before 2009. [b] Used to select the sample. [c] Unemployment size is the number of hours per week for which an individual is unemployed, based on the number of weekly working hours in the lost job.

Figure 1: Number of UI entrants per month

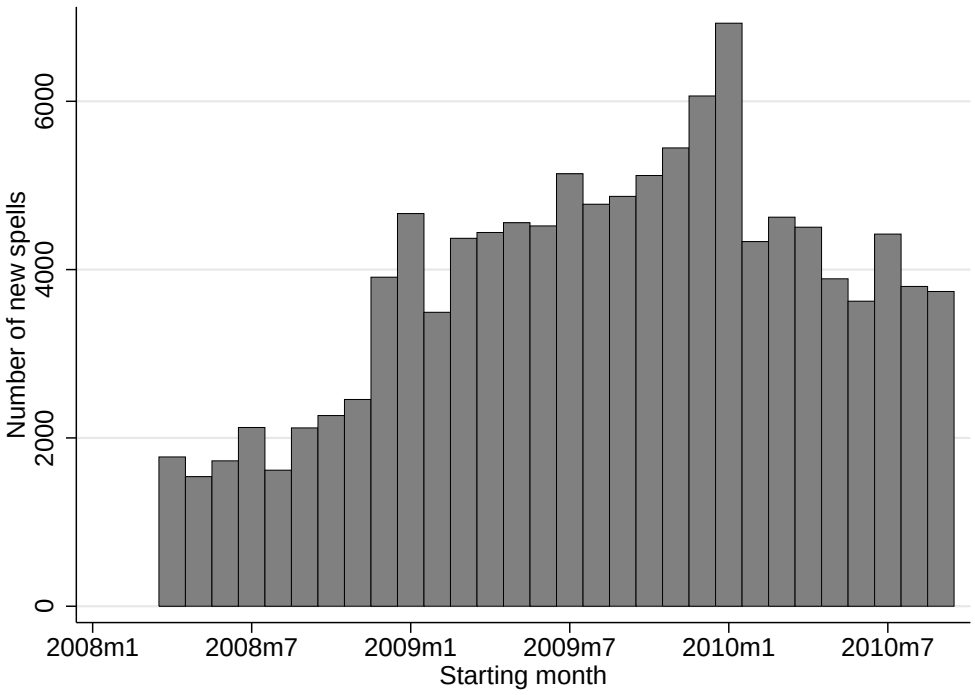
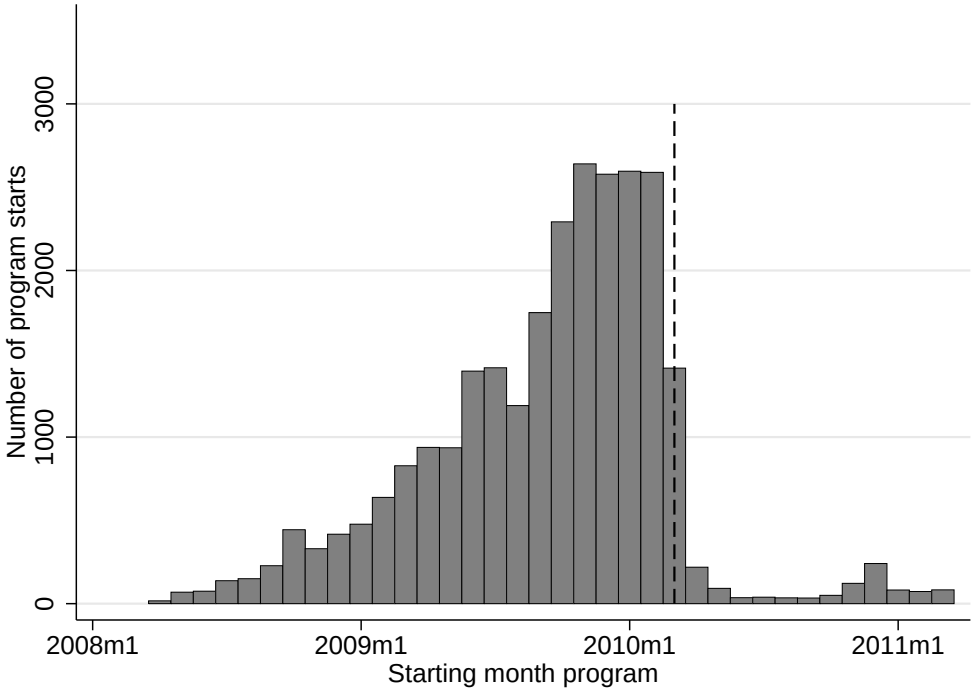


Figure 2: Number of UI benefits recipients entering an external program per month



to sickness or disability are unlikely to be affected by program participation (and are in any case not outcomes of interest). Therefore, we focus on exits to work and exits due to unknown reasons (these may often include situations in which workers generate some income from other sources than employment).

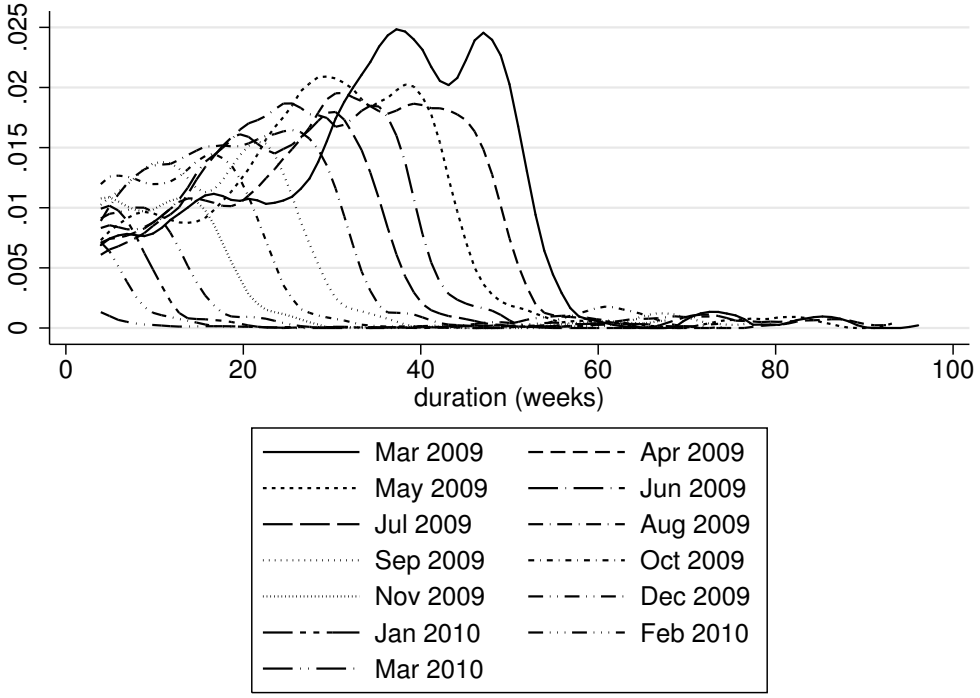
In the full sample about 19% of the benefits recipients participate in one of the externally provided programs. Two-thirds of these programs focus on job-search assistance, a third involve some sort of training, while only a very small fraction are subsidized employment. About 37% of all individuals participate in an internal program, of which the majority is either some test (such as a competencies test) or job-search assistance.

The dataset contains a large set of individual characteristics, including gender, age, immigrant status, education level, previous hourly wage, unemployment size, occupation in previous job, unemployment history, region and industry. In the lower panel of Table 1 sample means are presented for some characteristics. The average individual is almost full-time unemployed (37.2 hours) and 29% have already experienced a period of unemployment in the three years before entering UI.

In columns (2), (3) and (4) the same statistics are presented for three subgroups of individuals entering unemployment in April 2008, April 2009 and April 2010, respectively. The impact of the policy discontinuity in March 2010 becomes clear from the share of the April 2010 group that participate in an external program. It drops to almost zero. To illustrate the impact of the discontinuity in March 2010, we show the number of external programs started per month in Figure 2. The dashed line indicates the moment of the policy change in March 2010. The number of program entrants drops to almost zero in April 2010. Separate graphs for each type of program are included in the appendix (Figure 17) and show that the discontinuity occurs for all types of program.

The calendar date of entry in UI determines how the policy discontinuity affects individuals. Figure 3 shows for the different inflow cohorts the weekly probability

Figure 3: Hazard rate into the external programs by month of inflow

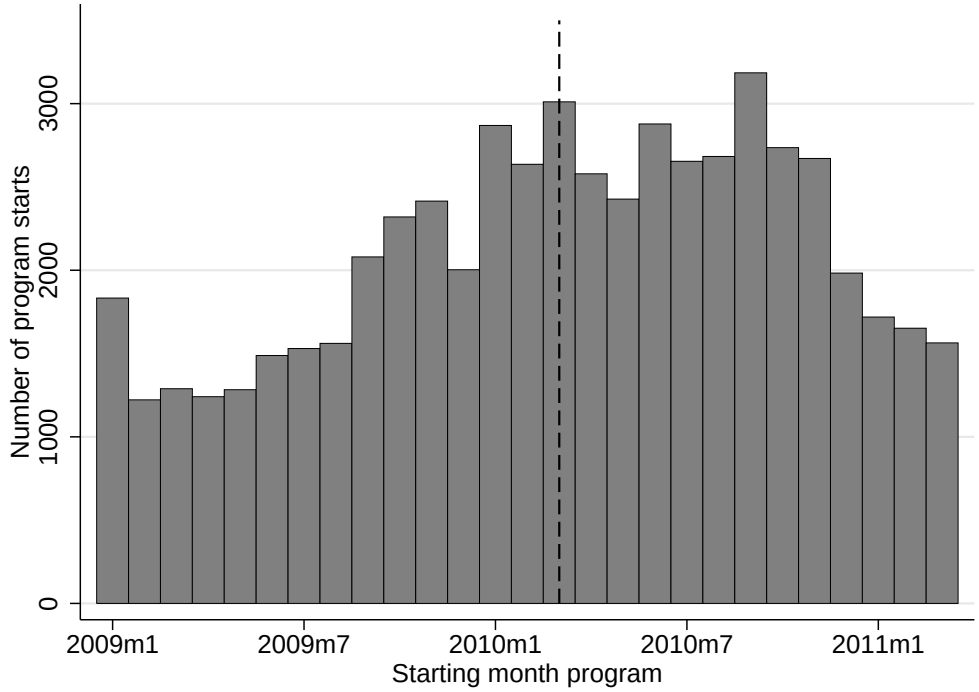


of starting an external program.[9] Each cohort reaches the policy discontinuity at a different moment in their UI spell. This is illustrated by the fact that each subsequent cohort experiences the drop in the program entry hazard one month earlier in their unemployment duration. The cohort of March 2010 has a probability of entering an external program close to zero. Figure 3 also shows that participation in some program is, in general, not restricted to a certain duration, though the hazard is increasing during the unemployment spell. Before the policy discontinuity the hazards of the different cohorts are very similar, indicating that there have not been other major policy changes.

A concern might be that caseworkers have responded to the inability to assign unemployed workers to external programs. However, resources for the internal programs remained unaffected around March 2010, limiting the scope for scaling up

[9]The graph shows the smoothed estimated hazard rate into the first external program of each individual.

Figure 4: Distribution of starting dates of the internally provided programs

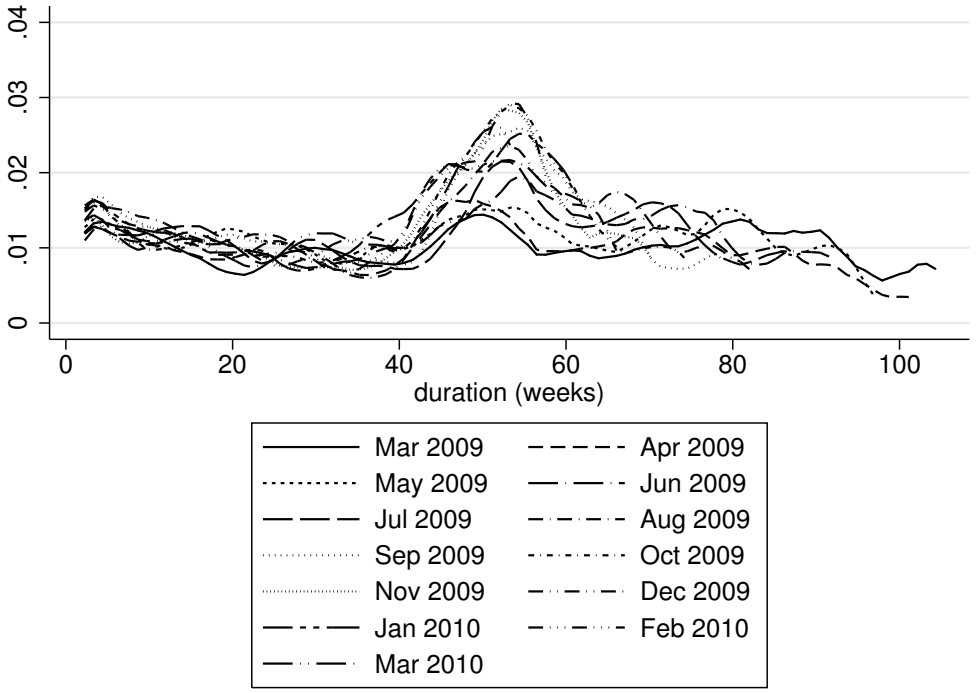


internal programs. The number of started internal programs per month is shown in Figure 4. Internal programs are only recorded from 2009 onward. The policy discontinuity of March 2010 is indicated by the dashed line. There is no indication of a response around the date of the policy discontinuity. Separate graphs by type of program are provided in the appendix (Figure 19). The hazard rates into an internal program for different cohorts are shown in Figure 5. The hazard rates are very similar, supporting the assumption that internal program provision was unaffected by the policy change.[10]

A further concern might be that even though the number of internal programs was not changed, caseworkers may have reacted to the unavailability of external programs by shifting their internal programs to these individuals that might otherwise have participated in external programs. This would imply that the policy

[10]In theory, job seekers could decide to pay for an external program themselves once it is no longer offered through the UI administration. However, the costs of these programs are considerable, especially for unemployed, such that this never happens in the Netherlands.

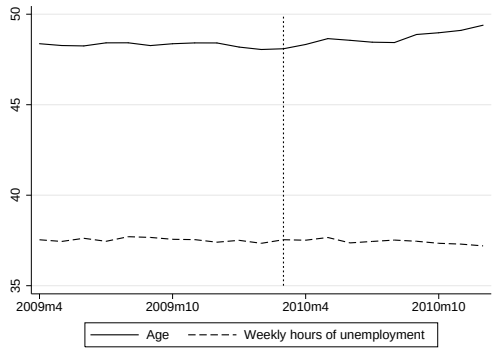
Figure 5: Hazard rate into the internal programs by month of inflow



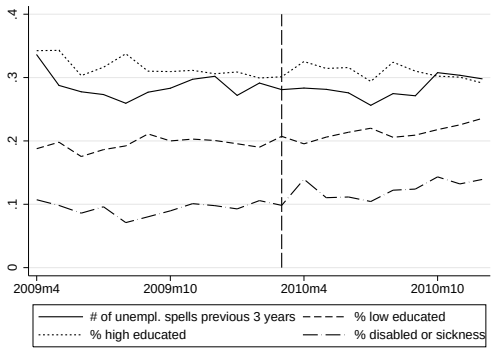
does not change external program participation to no participation, but, for some individuals, changes it to internal program participation.

To investigate whether such a shift of internal program targeting indeed occurred, we present sample means of characteristics of individuals enrolling in an internal program per month in Figure 6. Mean age and weekly hours of unemployment are shown in panel (a) of Figure 6, unemployment and disability history and education level are shown in panel (b), the previous hourly wage is shown in panel (c) and the share of nine industry categories is shown in panel (d). None of the graphs indicate any kind of discontinuity around March 2010, which suggests that caseworkers did not shift internal programs to individuals who would otherwise enroll in an external program. The effect of enrollment in an external program should thus be interpreted as the effect conditional on the allocation of internal programs.

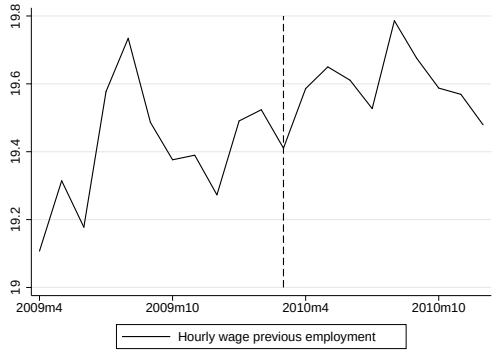
Figure 6: Composition of the internal program participants



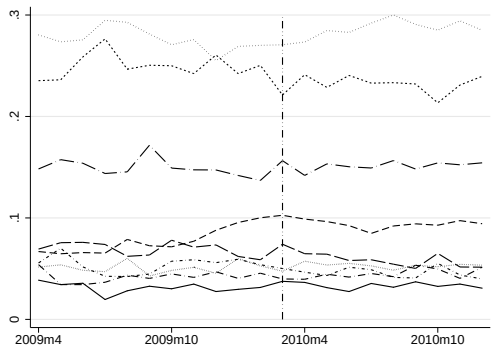
(a) Mean of age and hours unemployed by month



(b) Mean of characteristics by month



(c) Mean of hourly wage by month



(d) Mean of industry shares by month

5 Treatment effects

In this section we define the treatment effects that we aim to estimate. Recall that only a small share of all unemployed workers enter an external program during their unemployment spell. Due to the selectivity in the participation decision, the composition of program participants differs from non-participants. We focus on the average treatment effect on the treated (ATET). This treatment effect is nonstandard because enrollment in the program is dynamic.

Our key outcome of interest is duration until employment, which is a random variable denoted by $T > 0$. Define $Y_t = \mathbb{1}(T > t)$, a variable equal to one if the individual is still unemployed in period t, and zero otherwise. And define S to be a random variable denoting the duration at which program participation starts (with realized value s). Potential unemployment durations depend on program

participation. The dynamic nature of duration data implies that even for a single program many different treatment effects arise. The program can start at different durations, while the effect can be measured at different points in time (see for an extensive discussion of dynamic treatment effects Abbring and Heckman (2007)). We define potential outcomes when treated as:

$$Y^*_{1,t}(s) = \begin{cases} 1 & \text{if } T^*(s) > t \\ 0 & \text{if } T^*(s) < t \end{cases}$$

The potential outcome under no treatment is defined as $Y^*_{0,t} = \lim_{s\to\infty} Y^*_{1,t}$. We adopt the so-called no-anticipation assumption (Abbring and Van den Berg (2003)). This assumption imposes that program participation at duration s only affects potential outcomes at durations $t > s$. This allows us to write the potential untreated outcomes as

$$Y^*_{0,t} = Y^*_{1,t}(s) \qquad \forall s > t$$

The no-anticipation assumption is strict since it rules out that individuals can anticipate program participation prior to s by changing their job-search behavior accordingly. This is unlikely for the programs we discuss in this paper. Programs are assigned by caseworkers on an individual basis. There are no strict criteria for participation and each period only a small fraction of the unemployed workers can enroll, so it is impossible for job seekers to know in advance when they will enter the program.

Individuals leave unemployment after different durations, such that the composition of the survivors changes over time. This dynamic selection requires defining the subgroup for which the treatment effect is evaluated (see Van den Berg et al. (2014)). We are interested in individuals that received job-search assistance after s periods of unemployment. This implies that we should condition on surviving in

unemployment for s periods, so we consider the sub-population with $T > s$. We are interested in job finding for this group before $t > s$ periods. This effect is defined by Van den Berg et al. (2014) as the average treatment effect on the treated survivors ($ATTS(s,t)$), with equals

$$ATTS(s,t) = E\Big[Y^*_{1,t}(s) - Y^*_{0,t}\Big|T > s, S = s\Big] \tag{1}$$

This is the treatment effect that we focus on in the analysis.

5.1 Choice of samples

We compare estimates for $ATTS(s,t)$ using different methods. The different methods allow or require samples that are not necessarily the same. Exploiting the policy discontinuity requires using a specific sample of individuals entering unemployment around the time of the discontinuity. Matching and timing-of-events can use a much larger sample including individuals entering unemployment earlier or later. We are interested in comparing how the conclusion regarding effectiveness of the programs depends on the method. Therefore, we argue that each method should be applied 'optimally', that is, as it would have been applied in a stand-alone evaluation. As a result, we use the larger available sample when the method allows, and the smaller specific sample otherwise.

Any resulting difference in results can be due to different samples. We argue

Table 2: Methods and samples

	(1) Full sample: Inflow between April '08 and Sept. '10	(2) Discontinuity sample: Inflow between Oct. '09 and Jan. '10	(3) Pre-disc. sample[a] Inflow between April '08 and Jan. '10
Matching	yes	yes	yes
Timing-of-events	yes	yes	yes
Quasi-experiment	no	yes	no

[a] In addition to restricting the inflow period, this sample also censors observations at the discontinuity.

that the sample selection is an essential part of the method. However, to investigate to what extent the sample choice drives the results, we perform each analysis also with a smaller sample that is the same across all methods (column (2) in Table 2). A more extensive discussion on the selection of the smaller sample is presented in Section 7, where we discuss the quasi-experimental approach exploiting the policy discontinuity. In addition, we also apply the matching and timing of events approaches to a third sample that excludes the discontinuity period (column (3) in Table 2). This sample contains only individuals entering unemployment before the discontinuity and censor all observations at the time of the discontinuity. The rationale for applying such a sample is that the discontinuity creates exogenous variation in program participation, and we study how non-experimental methods perform without including such variation.

Our comparison considers the approaches presented in Table 2. The full sample and the pre-discontinuity sample are used for the non-experimental methods only, while the discontinuity sample is used for all three methods.

6 Non-experimental analysis

6.1 Matching estimator

We start the empirical analysis by applying a matching estimator as is commonly used in the literature. This does not exploit the policy discontinuity, but instead compares individuals with similar characteristics differing only in treatment status. We apply a dynamic version of the matching estimator to account for the dynamic setting and selection.

Matching methods construct counterfactual outcomes in the absence of randomized assignment (Imbens and Wooldridge (2009), Abadie and Imbens (2011)) and date back to Rosenbaum and Rubin (1983). The approach relies on two main as-

sumptions. First, selection into treatment is on observables only:

$$Y^*_{0,t}(s), Y^*_{1,t}(s) \quad \perp S|X \tag{2}$$

This unconfoundedness assumption implies that after conditioning on a set of observed characteristics, assignment to treatment is independent of the potential outcomes.[11] Our administrative data include a rich set of covariates, which is crucial for matching estimators. Employment histories are argued to be particularly important, because they are strong predictors of future labor-market outcomes as well as program participation (see, for example, Card and Sullivan (1988), Heckman et al. (1999), Gerfin and Lechner (2002), Lechner et al. (2011) and Lechner and Wunsch (2013)). In addition to employment history (previous hourly wage, unemployment history, industry), we observe individual characteristics (age, gender, education level, marital status, region) and variables describing the current unemployment spell (unemployment size in hours, sickness or disability, maximum benefits entitlement). This set of covariates is as least as extensive as usually available when evaluating active labor-market programs.

Second, the matching estimator requires a common support in the distribution of the covariates between program participants and non-participants. For our dynamic setting we assume

$$f_S(s;x) > 0 \qquad \forall x, s$$

where $f_S(s;x)$ is the density function of enrolling in the program after s periods of unemployment conditional on the set of covariates x. At any duration, all individuals have a positive probability of starting treatment regardless of their characteristics. This ensures that if the sample size is sufficiently large, counterfactuals can always be found. This assumption is likely to hold, since there are no (combinations of) individual characteristics that perfectly predict program participation in our data.[12]

[11] Vikström (2016) for a discussion considering also a more general setting of the dynamic treatment evaluation based on selection on observables.

[12] Figure 23 in the appendix shows the predicted probabilities of ever enrolling in the program

Our baseline estimates are based on the full sample, and as robustness checks we use two restricted samples (see subsection 5.1). We focus on starting the job-search assistance program after three to five months. Since we are interested in the average treatment effect for the treated survivors, we condition on surviving in unemployment for at least three month. Furthermore, we censor unemployment duration for individuals who enter the program after five months unemployment. This approach is similar to Lalive et al. (2008) (see also for a discussion Sianesi (2004)).

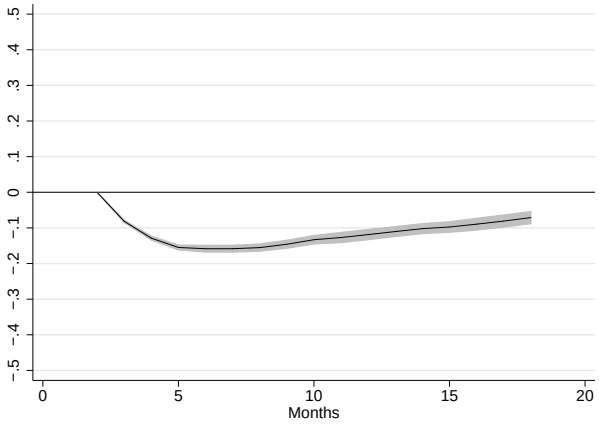
We use a logit model to predict program participation rates for all individuals surviving for at least three months in unemployment. Next, we estimate for both the treated and the control group the Kaplan-Meier estimates (taking censoring and future program participation into account) for survival in unemployment, where we weigh individuals in the control group to make the composition comparable to the treatment group.[13] The difference in the survival functions provides an estimate of the average treatment effect on the treated survivors.[14] This is shown in panel (a) of Figure 7 together with the 95%-confidence interval which is obtained using bootstrapping. Program participation significantly reduces the job finding probability, although the negative effect becomes smaller over time. This is consistent with a lock-in effect. The estimates state that those who participate in the job-search assistance program are about 20%-point less likely to have found work within five months after becoming unemployed. This effect reduces to about 7%-points 18 months after becoming unemployed. Due to the large sample size the confidence interval is very tight. Panel (b) and (c) of Figure 7 show that the estimated effects are not sensitive

(based on a logit model). This shows a large overlap which is usually considered as support for the common support assumption (Busso et al. (2009)). The latter is due to the dynamic selection in our setting, not true (i.e. individuals with particular characteristics may not receive treatment because they leave unemployment quickly). However, when conditioning on survivors, we observed a similar overlapping support.

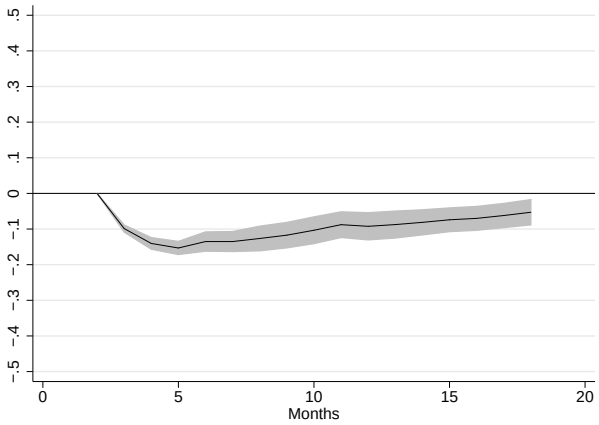
[13]Individuals with characteristics x_i in the control group get weight $\hat{p}(x_i)/(1-\hat{p}(x_i))$, where $\hat{p}(x_i)$ is the predicted treatment participation probability.

[14]Our approach is similar to the inverse probability weighting estimator described by Vikström (2016), which also takes future program participation (for the control group) en right-censoring of unemployment durations into account.

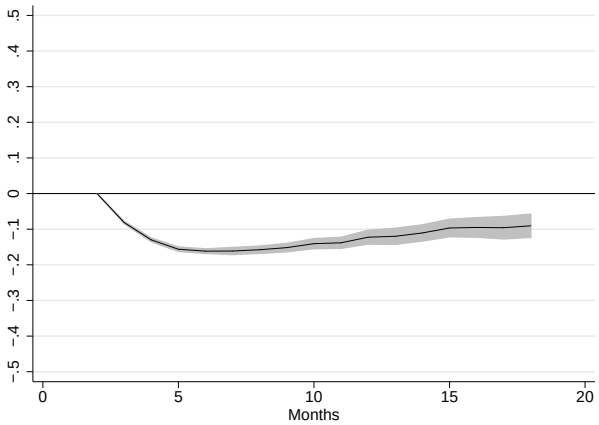
Figure 7: Average treatment effect on the treated survivors (with 95% confidence intervals): Matching estimator using Kaplan-Meier survival functions



(a) Inflow between April 2008 - September 2010



(b) Inflow between October 2009 - January 2010



(c) Inflow between April 2008 - January 2010 (and censored by March 2010)

to the choice of the observation period. In both cases confidence intervals are wider because reduced samples are used for the estimation.

6.2 Timing of events model

Matching requires little functional-form assumptions, but relies on a potentially strong unconfoundedness assumption. The timing-of-events model (Abbring and Van den Berg (2003)) allows for selection on unobservables, but makes stronger functional-form assumptions. This model has been applied often in the recent literature on dynamic treatment evaluation (see for example Abbring et al. (2005), Van den Berg et al. (2004), Lalive et al. (2005) and Van der Klaauw and Van Ours (2013)).

The timing-of-events model jointly specifies job finding and entry into the program using continuous-time duration models. To control for unobserved characteristics the unobserved heterogeneity terms in both hazard rates are allowed to be correlated. Identification relies on the mixed proportional structure of both hazard rates. As discussed in subsection 7.1, the timing-of-events model requires the no-anticipation assumption (as do the quasi-experimental approach and the matching estimator). Note that this does not rule out that the treatment probability differs between individuals and that individuals are aware of this. Some job seekers may have a high probability of program assignment and know this. Only the exact timing of the program start should be unanticipated.

We present a concise description of the model here, while a detailed version is presented in the appendix. Consider an individual entering unemployment at calendar date τ_0. The job finding (hazard) rate depends on the number of days of unemployment t, calendar time $\tau_0 + t$, observed characteristics x and unobserved characteristics v_e. When starting the job-search program after s periods of unemployment, the hazard rate shifts by the treatment effect δ_{t-s}, which can depend on the elapsed duration $t - s$ since entering the program. The treatment effect is

modeled as piecewise constant function of the elapsed duration since starting the program (see Appendix A for the parameterization). The job finding rate is given by:

$$\theta_e(t|x, \tau_0, s, v_e) = \phi_e(t)\psi_e(\tau_0 + t) \exp\Big[x\beta_e + \delta_{t-s} I(t > s)\Big] v_e \tag{3}$$

Estimation of equation (3) yields a biased estimate of the treatment effects if program participation is (even conditional on the observed characteristics) non-random. To account for this, program participation is modeled jointly, also using a mixed proportional hazard rate:

$$\theta_p(s|x, \tau_0, v_p) \quad = \quad \phi_p(s)\psi_p(\tau_0 + s) \exp(x\beta_p) v_p \tag{4}$$

With all notation similar to equation (3), but subscript e replaced by subscript p. The unobserved term v_p is allowed to be correlated with v_e, with joint discrete distribution $g(v_e, v_p)$. We take $g(v_e, v_p)$ to be a bivariate discrete distribution with an unrestricted number of mass points. The duration dependence patterns and the calendar time effects are parameterized with a piecewise constant function. Estimation of the parameters is performed by maximizing the log-likelihood, in which right-censoring is straightforwardly taken into account.

The model is estimated using the full sample, as well as using the smaller discontinuity sample and the pre-discontinuity sample. Full estimation results, including all estimated coefficients, are presented in the appendix in Table 4.[15]

The estimates of the program participation effects are presented in the first column of Table 4. The effect of program participation is estimated to have a large, significantly negative effect on the job finding rate in the first three months ($\delta_{0-3 \text{ months}}$), with the hazard ratio equal to 0.723. In the next three months the effect is still significantly negative, but smaller in magnitude (0.873). After six

[15]The coefficients of the job finding hazard (equation (3)) have the expected sign and most are highly significant. Outflow from unemployment is higher for higher education levels, higher income and lower age. Having a sickness or disability history decreases outflow, while being married or cohabiting increases outflow. Furthermore, we find that the job finding rate decreases steadily over unemployment duration.

Table 3: Treatment effect estimates Timing-of-Events model

	(1) Full sample		(2) Discontinuity sample		(3) Pre-discont. sample	
	Coef.	st.er.	Coef.	st.er.	Coef.	st.er.
Program effect on UI outflow:						
$\delta_{(1-3 \text{ months})}$	0.723	0.021	0.765	0.043	0.606	0.032
$\delta_{(4-6 \text{ months})}$	0.873	0.022	0.796	0.047	0.738	0.042
$\delta_{(>6 \text{ months})}$	1.080	0.016	1.046	0.036	0.853	0.056
Observations	116,625		23,502		83,773	

Reported values are hazard ratios. The full sample contains all individuals entering unemployment between April 2008 and September 2010. The discontinuity sample contains all individuals entering unemployment between October 2009 and January 2010. The pre-discontinuity sample contains all individuals entering unemployment between April 2008 and January 2010, and censors all observations at the time of the discontinuity (March 2010).

months ($\delta_{\geq 6 \text{ months}}$) program participation has a modest but significantly positive effect on the probability of finding a job (1.080). When using the smaller "discontinuity" sample of individuals entering unemployment in October 2009 - January 2010, we find very similar estimates for the program effects: a negative effect over the first six months, and a positive effect afterwards (see column (2) in Table 3). Standard errors are larger due to the smaller sample size. The third sample on which we estimate the model is the pre-discontinuity sample. Note that by excluding observations from the discontinuity period, we exclude all exogenous variation in program participation. Results are presented in column (3), and we find somewhat more negative effects on outflow. Furthermore, the negative impact on outflow remains even after six months.

In the first two estimations we find no evidence of unobserved heterogeneity. That is, the probability of all mass points except one converge to zero. This finding suggests that, when conditioning on all observables in our dataset, little additional factors remain that determine both outflow from unemployment and program assignment. However, when using the third sample, we do find some unobserved

heterogeneity. The fact that the third sample leads to more negative estimates, suggests that there is negative selection. Those that participate in programs have, on average, worse labor market prospects. The first two samples, that include variation in participation due to the discontinuity, suffer from this problem to a lesser extent. However, the relatively small difference in estimates suggests that the selection problem is modest.

The estimates for the parameters δ provide a multiplicative effect on the job finding rates, but can not be directly interpreted as measure for the treatment effects $ATTS(s,t)$. Therefore, we follow Kastoryano and Van der Klaauw (2011), who define for unemployed worker i with observed characteristics x_i

$$\mathrm{E}[Y^*_{1,t}(s) - Y^*_{0,t} | T > s; x_i, v_e] = \frac{\exp(-\int_0^t \theta_e(z|x_i,t,v_e)dz) - \exp(-\int_0^t \theta_e(z|x_i,s,v_e)dz)}{\exp(-\int_0^s \theta_e(z|x_i,s,v_e)dz)}$$

To translate this in the average treatment effect on the treated survivors, we should condition on the rate of receiving treatment after s periods. Therefore, we use the hazard rate model for entering the program, which gives

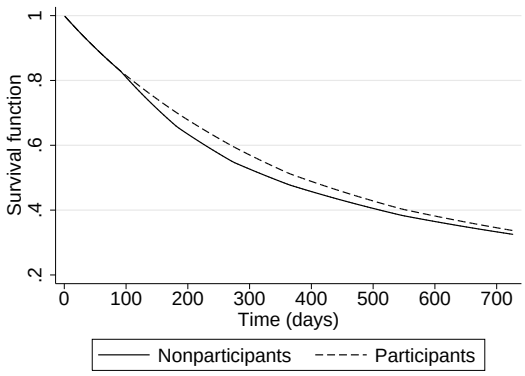
$$ATTS(s,t) = \frac{\int_v \sum_i f(s|x_i,v_p)\mathrm{E}[Y^*_{1,t}(s) - Y^*_{0,t}|T > s; x_i, v_e]dG(v_e,v_p)}{\int_v \sum_i f(s|x_i,v_p)dG(v_e,v_p)} \tag{5}$$

where $f(s|x_i,v_e,v_p) = \theta_p(s|x_i,v_p)\exp(-\int_0^s \theta_e(z|x_i,v_e,s) + \theta_p(z|x_i,v_p)dz)$ is the rate at which individual i enter the job search assistance program after s periods. We use the delta method to compute standard errors around the treatment effects.

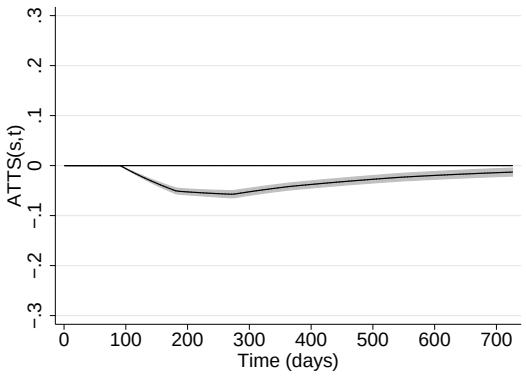
In panel (a) of Figure 8 we present simulated survivor functions for an untreated job seeker and for a job seeker starting a program after two months.[16] Participating in a program after two months of unemployment lowers the probability of being employed subsequently, in accordance with the negative effect estimate. In panel (b) we present the treatment effect (ATTS, equation (5)) with a 95% confidence

[16]These are for nonparticipants $\int_v \left[N^{-1}\sum_i \exp(-\int_0^t \theta_e(z|x_i,s,v_e)dz\right]dG(v_e,v_p))$ and of program participants $\int_v \left[N^{-1}\sum_i \exp(-\int_0^t \theta_e(z|x_i,t,v_e)dz)\right]dG(v_e,v_p)$.

Figure 8: Timing-of-events model: treatment effect with full sample



(a) Survivor functions



(b) Treatment effect (with 95% confidence interval)

interval computed using the delta method. The difference is significantly negative directly after the program starts, and increases in magnitude up to almost 6%-points after six months. At longer durations the difference decreases in magnitude and converges to zero.

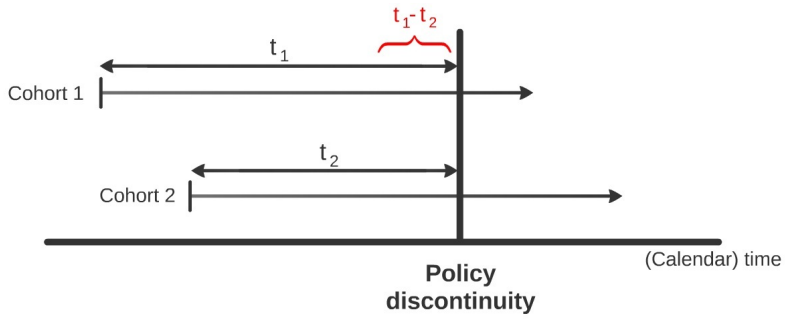
7 Quasi-experimental analysis

We now turn to the policy discontinuity to investigate how exploiting the exogenous variation compares to the estimates from the previous section. We start by characterizing how the discontinuity allows identification of the effect of the program on outflow to work.

7.1 Identification

Estimating the ATTS (equation (1)) without making parametric assumptions is generally not possible from observational data (Abbring and Van den Berg (2005)). A policy discontinuity provides exogenous variation which allows to estimate the ATTS (Van den Berg et al. (2014). Consider two cohorts of entrants in unemployment. The first enters unemployment at some point in time, with the time until the policy change equal to t_1. The second cohort enters unemployment later, but still before

Figure 9: Treatment effect identification



the policy change. For this cohort the time until the policy change equals $t_2 < t_1$. This is illustrated in Figure 9. The two cohorts face the same policy of potential program assignment for t_2 time periods, implying that dynamic selection is the same up to this point. After t_2, the first cohort faces another period of potential program assignment, with length $t_1 - t_2$, while the second cohort is excluded from program participation. As a result, we can compare the outflow to employment in the two cohorts, for those individuals that survived up to t_2 and did not enroll in a program prior to t_2.

To assign differences in job finding between these groups to program participation, two assumptions should hold. First, the policy discontinuity should be unanticipated by unemployed workers and caseworkers. Anticipation of the policy discontinuity is problematic, as behavior of job seekers or caseworkers may be affected in the period just before March 2010. The policy change we are investigating however has the advantage that it was unexpected. The UI administration only realized late that there was no longer budget for these programs and expected that the Ministry of Social Affairs would extend the budget. Enrollment in the external programs stopped immediately after an extension of the budget was rejected. Since not even the UI administration was expecting the change, we can safely assume that job seekers and caseworkers did not anticipate it either. The second assumption is that there should be no differences between the two cohorts in factors that

affect job finding, other than the difference in program assignment. We discuss this assumption below.

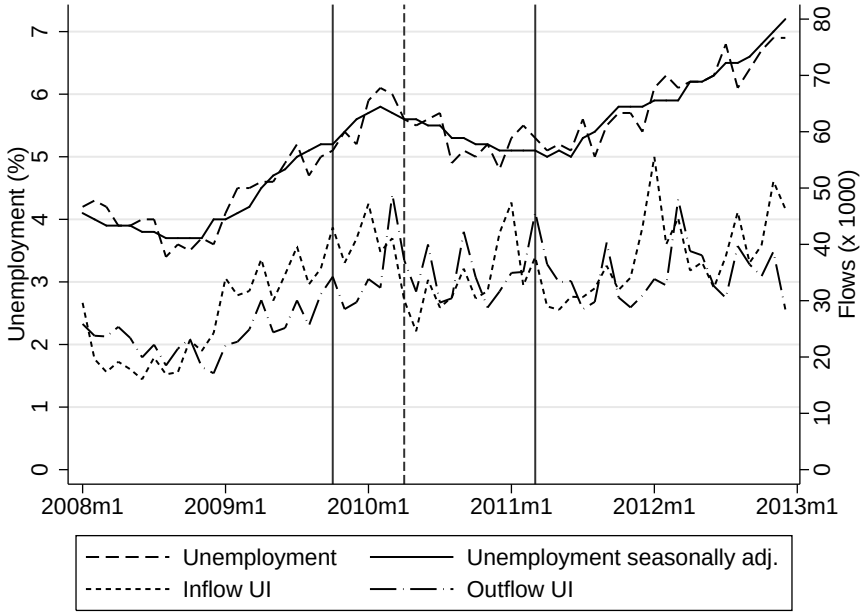
7.1.1 Business cycle, seasonalities and cohort composition

Even if two cohorts are compared that enter unemployment relatively shortly after each other, changes in the labor market conditions may lead to differences in outcomes. We discuss how this may affect our estimates, and how we correct for this. Figure 10 presents the unemployment rate and the inflow and outflow of unemployment. The two vertical full lines indicate the observation period that is used in the analysis. The vertical dashed line indicates the policy discontinuity. In the period before the policy discontinuity, 2009 and the beginning of 2010, unemployment was rising. During 2010 it decreased slightly, while in 2011 it started increasing again. In the short-run, seasonalities are the main source of fluctuations in unemployment. Also inflow into and outflow from UI are relatively stable around the policy discontinuity, except for short-run fluctuations. Such fluctuations in labor market conditions may affect outcomes in two ways. First, they affect the composition of the inflow into unemployment. For example, the financial crisis may cause that different types of workers become unemployed. A changing composition affects aggregate outflow probabilities. Second, labor market conditions affect outflow probabilities directly, as it is often more difficult to find employment when unemployment is high.

To correct for differences in composition we exploit the set of covariates in the data. In particular, we use weights to make each cohort in composition of observed characteristics comparable to the March 2010 cohort. As characteristics we use three previous hourly wage categories, an indicator for having been unemployed in the past three years, an indicator for being married or cohabiting, age categories, an indicator for being part-time unemployed (less than 34 hours per week)[17] and three education categories. Interacting these covariates we obtain 288 groups. Define the share of group g in cohort c by $\alpha_{c,g}$. The weight assigned to an observation belonging

[17]Since this is based on the previous job, it captures part-time and full-time employment.

Figure 10: Labor market indicators



Source: Centraal Bureau voor de Statistiek (CBS), Statline.

to group g in cohort c is given by:

$$w_{c,g} = \frac{\alpha_{mar2010,g}}{\alpha_{c,g}}$$

We define the survivor functions that will be estimated in the analysis, as the weighted average of the survivor functions of each cohort-group:

$$\bar{F}_c(t) = \sum_g w_{c,g} \bar{F}_{c,g}(t) \tag{6}$$

These weights are applied in all further analysis, however the results are robust against using weights.

The direct effect of the business cycle and seasonal effects on employment probabilities requires further discussion. To formalize these factors, consider the following simple model. Assume that the hazard rate to employment (h) for cohort c depends on the duration of unemployment (t), the effect of the business cycle (b_c), the effect of seasonalities (l_c) and the effect of entry into a program at time s, which is $\gamma(s)$.

To correct for business cycle effects when identifying the effect of program participation, we need to make some assumptions about the hazard. We assume that the business cycle, seasonalities and treatment have an additive effect on the baseline hazard, where each of these impacts may vary by unemployment duration t. Note that this is very flexible as we do not assume anything on how these factors vary by duration. The duration dependence of the hazard is denoted by $\lambda(t)$, which is common for all cohorts. The hazard rate is given by

$$h(t,s,c) = \lambda(t) + b_c(t) + l_c(t) + \gamma(s,t) \tag{7}$$

From the hazard rate we can construct the survival function.

$$\begin{aligned}
\bar{F}_c(t) &= \exp\big(-\int_0^t h(u)du\big) \\
&= \exp\big(-\int_0^t \lambda(u)du - \int_0^t b_c(u)du - \int_0^t l_c(u)du - \int_s^t \gamma(u)du\big)
\end{aligned}$$

Taking the logarithm of the survival function we have:

$$\begin{aligned}
\log \bar{F}_c(t) &= -\int_0^t \lambda(u)du - \int_0^t b_c(u)du - \int_0^t l_c(u)du - \int_s^t \gamma(u)du \\
&\equiv \Delta(t) + B_c(t) + L_c(t) + \Gamma(s,t)
\end{aligned}$$

This implies that the business cycle, seasonalities and program participation have additive impacts on the log of the survival function. Seasonalities are by definition those factors that are common across different years, such that $L_c(t) = L_{c-12}(t)$, for all c. The difference in log survivor functions of two cohorts identifies the treatment effect plus the difference in seasonal and business cycle effects. For example, comparing the January 2010 cohort with the October 2009 cohort we get:

$$\begin{aligned}
\mu^1(t) &= \log \bar{F}_{jan10}(t) - \log \bar{F}_{oct09}(t) \\
&= \Gamma(s,t) + \big[L_{jan10}(t) - L_{oct09}(t)\big] + \big[B_{jan10}(t) - B_{oct09}(t)\big]
\end{aligned} \tag{8}$$

If we condition both survivor functions on survival up to t_2 (in this case t_2 = two months), the January 2010 cohort never enters a program, while a share of the October 2009 cohort enters a program. The term $\Gamma(s,t)$ measures the effect of program participation of a share of a cohort, and can thus be interpreted as an intention-to-treat effect. Below we discuss how to this relates to the ATTS (equation (1)). The size of the bias due to the remaining terms, $(L_{jan10}(t) - L_{oct09}(t))$ and $(B_{jan10}(t) - B_{oct09}(t))$, depends on the length of the time interval between the two cohorts, and the volatility of the labor market.

We can possibly improve on this estimator by applying an approach related to a difference-in-differences estimator. By subtracting the same cohort difference from a year earlier, we eliminate the seasonal effects, at the cost of adding extra business cycle effects:

$$\begin{aligned} \mu^2(t) &= \left[\log \bar{F}_{jan10}(t) - \log \bar{F}_{oct09}(t)\right] - \left[\log \bar{F}_{jan09}(t) - \log \bar{F}_{oct08}(t)\right] \\ &= \Gamma(t) + \Big[B_{jan10}(t) - B_{oct09}(t)\Big] + \Big[B_{jan09}(t) - B_{oct08}(t)\Big] \end{aligned} \tag{9}$$

Whether this is preferable over $\mu^1(t)$ depends on the relative sizes of the business cycle and seasonal effects. Figure 10 suggests that, if the interval is sufficiently small, seasonal effects are much larger than business cycle effects. Given a small interval such as three months, business cycle effects may be small enough to ignore, such that μ^2 is a satisfactory estimator of $\Gamma(t)$. Note that this estimator is an extension of the approach suggested by Van den Berg et al. (2014), who exploit a policy discontinuity to estimate effects on a duration variable. We add to this approach by taking double differences.

The estimators $\mu^1(t)$ and $\mu^2(t)$ estimate intention-to-treat effects, since not all unemployed workers in the earlier cohort enter a program. The average treatment effect on the treated survivors follows from dividing the intention to treat effect by the difference in the share of each cohorts that enrolls in the program. Define $\bar{F}^{treat}$ as the survivor function for treatment, where an exit is defined to be the start of

the first program. The ATTS estimator is given by:

$$ATTS(t_2, t) = \frac{\mu^1(t)}{\log \bar{F}^{Treat}_{jan10}(t) - \log \bar{F}^{Treat}_{oct09}(t)} \tag{10}$$

And similar for $\mu^2(t)$.

7.2 Results: intention-to-treat effect

We start by defining which cohorts to compare. A cohort contains all individuals entering unemployment within one particular month. The time interval between cohorts should be small to minimize business cycle and seasonal effects, but the trade-off is that more time between cohorts increases the difference in exposure to potential program participation. We use cohorts three months apart. Second, to exploit the policy discontinuity, the cohorts should not enter unemployment too long before March 2010. Therefore, we use the cohorts of October 2009 until January 2010, facing between five and two months of potential program participation, respectively. Each cohort will be compared to the cohort entering unemployment three months earlier. The survivor function of each cohort is presented in Figure 11. Around 50% of the UI benefits recipients find work within 12 months, while after two years around 65% has found work.

We first take the difference between the log of the survivor function and the log of the survivor function of the cohort entering unemployment three months earlier ($\mu^1(t)$).[18] This compares the outflow of a cohort from in which no one enrolls in the program to outflow of a cohort in which a share enrolls in the program. As discussed in subsection 7.1, we condition on survival and no-treatment up to the duration at which the later cohort reaches the policy discontinuity. So when comparing January 2010 with October 2009, only individuals are included with an unemployment duration of at least three months and who do not start an external

[18] All estimates presented in this section are estimated using weights as discussed in subsection 7.1.1.

Figure 11: Survivor functions by month of inflow

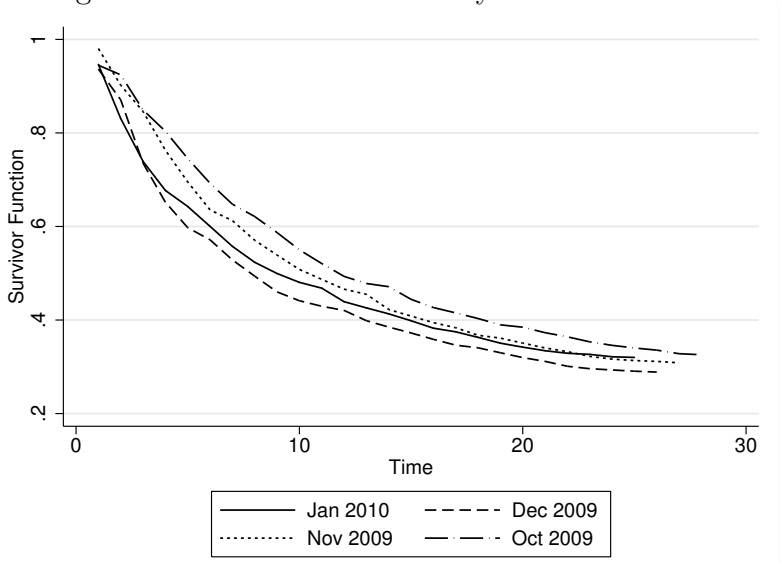
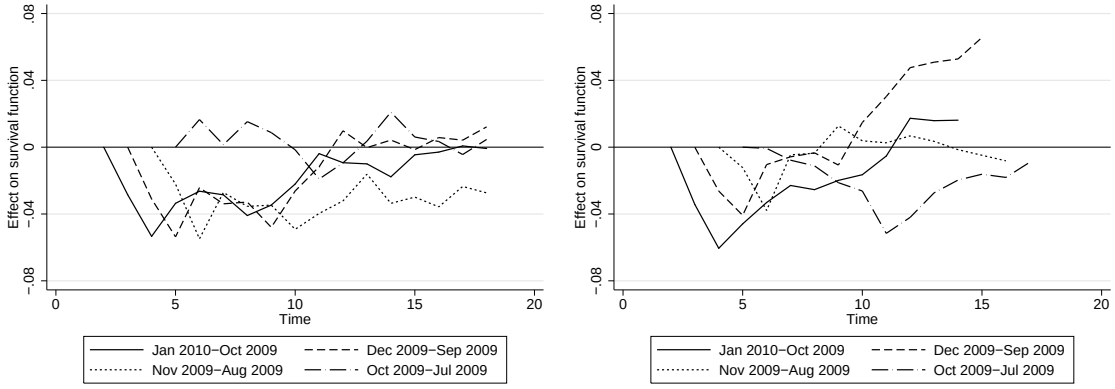


Figure 12: Intention-to-treat effect estimates, conditional on $T > t_2, S > t_2$



(a) $\mu^1(t)$ (b) $\mu^2(t)$

program in the first two months. The differences up to a duration of 18 months after inflow in unemployment are presented in panel (a) of Figure 12.[19] We find a negative effect on job finding during the first few months after program participation of around 4%-points in three of the four estimates. After about 10-12 months the negative effect disappears and all estimates are close to zero.

These estimates are based on simple differences between cohorts, thus not taking

[19]For the ease of interpretation, in the graph we present a transformation of the estimates μ_1. We display $[\exp(\mu_1) - 1]\bar{F}$, which is the effect on the actual survival function. So the graph can be interpreted as the effect on the probability of finding employment. The same transformation is made in subsequent graphs.

fluctuations in labor market conditions into account. By subtracting the same differences from a year earlier, we correct for cohort differences that are constant across years (such as seasonalities). Estimates from such a "difference-in-differences" approach ($\mu^2(t)$) are presented in panel (b) of Figure 12.[20] Again we find a negative effect on job finding in the first months. At longer durations, the estimates diverge somewhat. In the appendix we present for each line a 95% confidence interval (standard errors are computed using bootstrapping) in Figure 20. The early negative effect is always significantly different from zero, while none of the estimates at longer durations are significantly different from zero. Note that each comparison measures the effect of additional treatment at a slightly different duration. For example, the January 2010-October 2009 comparison measures the effect of additional treatment in the 3th-5th month of unemployment, while the December 2009-September 2009 comparison measures the effect of additional treatment in the 4th-6th month of unemployment.

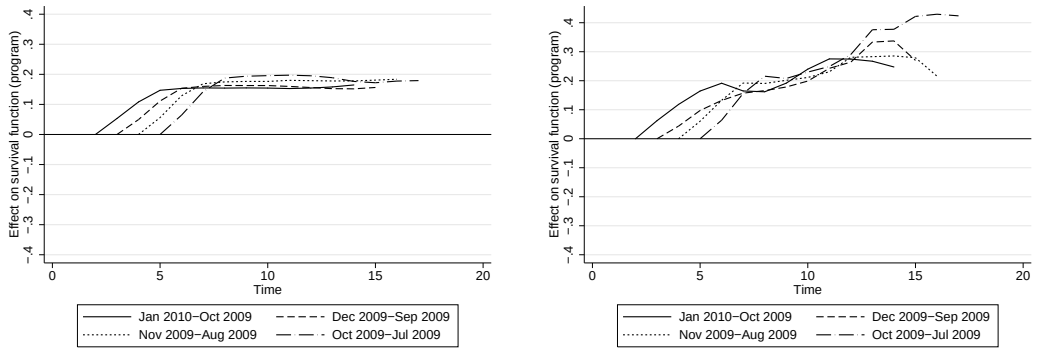
The results show a pattern that is quite consistent across different cohort comparisons and across the two estimators. Job finding is significantly reduced in the early months, while the difference disappears after 6-12 months. This finding is in line with the lock-in effect often found in the literature (see for example Lechner and Wunsch (2009)). The lock-in effect implies that when a program starts, participants shift attention from job search to the program which reduces their job finding rate. This negative effect disappears after some months, but we do not find any (positive) effects at longer durations.

7.3 Results: average treatment effect

The above findings are intention-to-treat effects. To estimate the average effect of treatment on individual employment probabilities they need to be scaled by the differences in treatment intensity. We divide each estimate by the difference in

[20] When estimating $\mu^2(t)$ we only present estimates up to the duration at which the cohorts from a year earlier reach the policy discontinuity, which is between 15 and 18 months. Estimates at longer durations are biased as the earlier cohorts are affected by the policy discontinuity as well.

Figure 13: Differences in program participation, conditional on $T > t_2$ and $S > t_2$



(a) Simple difference

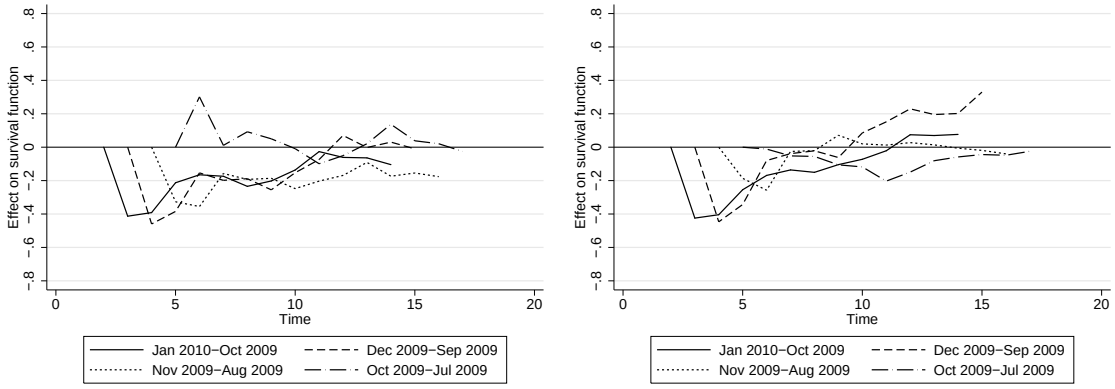
(b) Double difference

program participation of the cohorts that are being compared (as defined in equation (10)). The difference in program participation occurs during a three months period in which the later cohort reached the policy discontinuity but the earlier cohort did not. We estimate the difference in program participation by the difference between the survivor functions for program participation (as defined in subsection 7.1.1). For example, when comparing the January 2010 cohort with the October 2009 cohort, the difference in program participation is given by

$$\bar{F}^{treat}_{Jan10}(t|T, S > \text{ 2 months}) - \bar{F}^{treat}_{Oct09}(t|T, S > \text{ 2 months}) \tag{11}$$

And similar for the other cohorts that are compared. These estimates are presented in panel (a) of Figure 13. We find a clear increase as soon as the first cohort reaches March 2010. The difference increases for approximately three months, after which the comparison cohort reaches March 2010. From that point onward, both cohorts receive no treatment, and the difference remains stable. The difference is between 15 and 20%-points. The difference in program participation can also be computed using the same "difference-in-differences" approach as in $\mu^2(t)$. Such estimates are presented in panel (b) of Figure 13. Due to differencing with cohorts from a year earlier, the differences are less smooth, though the main pattern remains. Confidence intervals are presented in Figure 21 in the Appendix. All differences are highly

Figure 14: Average treatment effect, conditioned on $T > t_2, S > t_2$



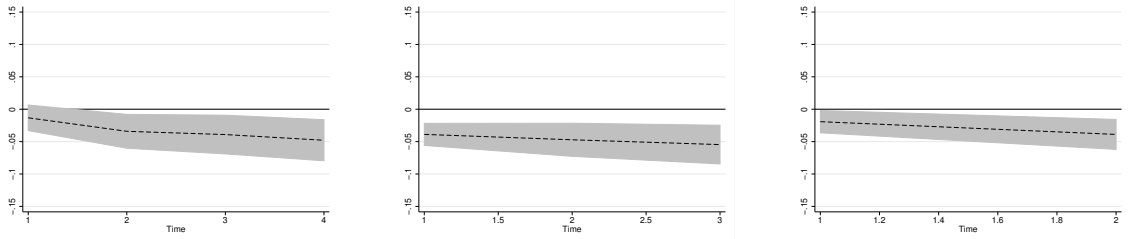
(a) based on $\mu^1(t)$ (b) based on $\mu^2(t)$

significant.

The results of dividing the estimates from panel (a) in Figure 12 by the treatment difference are presented in panel (a) of Figure 14. The pattern does not differ much from that of the intention-to-treat effects. Program participation reduces the job finding probability during the first two-three months by about 40%-point, while after ten months employment probabilities are similar again and there is no significant effect. The double differencing estimates ($\mu^2(t)$) are presented in panel (b) of Figure 14. The pattern is quite similar. There is a negative effect on the job finding probability of 40%-point directly after program participation starts, which decreases in magnitude over time towards a zero effect after about eight months (confidence interval are presented in Figure 22 in the Appendix).

Since the policy discontinuity reduced program participation to zero, these estimates can be interpreted as average treatment effects on the treated, rather than local average treatment effects. Alternatively, if the policy discontinuity only had reduced program participation (but not to zero), our approach would have estimated the treatment effect on those individuals that would have participated before the policy discontinuity but not afterwards. In the terminology of instrumental variables, these are the "compliers". The fact that the policy discontinuity reduced program participation to zero, implies that there are no "always-takers", and the

Figure 15: Common trend tests, $\mu^2(t)$, three months differences



(a) December 2009 - September 2009

(b) November 2009 - August 2009

(c) October 2009 - July 2009

local average treatment effect equals the average treatment effect.

7.4 Common trend assumption

The assumption that the business cycle terms in (9) are negligible has some similarities with the common trend assumption in a difference-in-differences estimator. It requires that in the absence of the policy discontinuity, the difference in employment rate between the January and November cohort would have been the same in 2009/2010 as a year earlier in 2008/2009. This is by definition not testable. However, we can get an indication of the plausibility of the assumption, by investigating the survivor functions over the first months of each cohort, so before exposure to the policy discontinuity. All estimators condition on survival up to t_2, but we can use information on job finding before t_2 to get some indication about the validity of our common trend assumption. To have a sufficient number of pre-discontinuity months in the latest cohort, we focus on the comparisons of December 2009, November 2009 and October 2009. Basically, we estimate $\mu^2(t)$ for $t \leq t_2$, without conditioning on survival up to a certain duration.

Estimates are presented in Figure 15 for December 2009 - September 2009, November 2009 - August 2009 and October 2009 - July 2009, including 95% confidence intervals, computed using bootstrapping. We find that all estimates are significantly negative, though small in magnitude. This implies that the treatment effect estimates are biased downwards somewhat, and thus provide lower bounds of

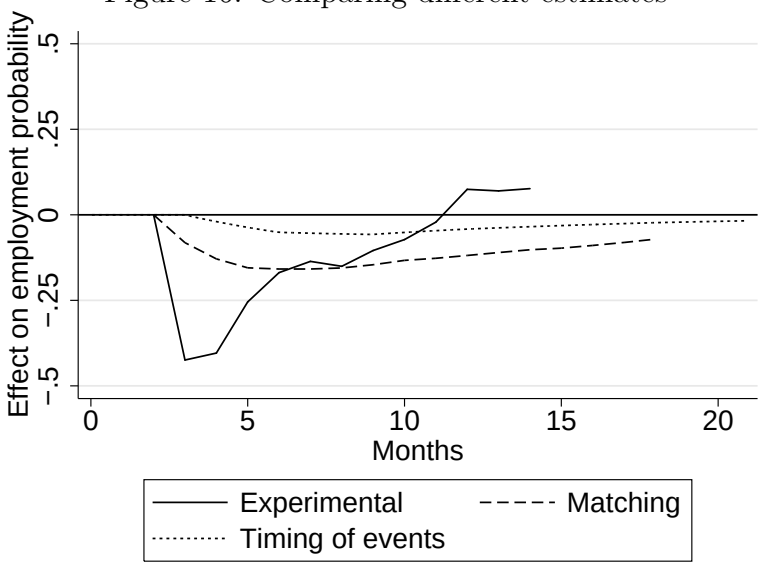
the effect.

8 Discussion

We have estimated the impact of the (external) activation programs on the exit rate to work using three approaches, (*i*) a dynamic matching estimator, (*ii*) the timing-of-events model, and (*iii*) exploiting a policy discontinuity as natural experiment. In this section we compare the results and discuss similarities and differences.

Figure 16 presents the three estimates in one graph.[21] The matching estimate corresponds to panel (a) in Figure 7. The timing-of-events estimate corresponds to the difference in simulated survivor functions presented in panel (b) of Figure 8. The quasi-experimental estimate is the one based on comparing the cohorts of January 2010 and October 2009 (panel (b) of Figure 14).[22] All three estimates represent the effect of participating in a program after three to six months of unemployment on the probability of finding work.

Figure 16: Comparing different estimates



[21] Note that for the matching estimates and the timing-of-events estimates we performed the analysis both using the full sample, and using the smaller "discontinuity" sample and the pre-discontinuity sample. Since the results are similar we focus on the estimates using the full sample.

[22] For simplicity we focus on this particular estimate, but the other estimates follow a similar pattern.

Each method measures the same treatment parameter and the estimates should thus be similar if all assumptions of the estimators are valid. Indeed, We find that the three methods lead to a similar conclusion. Immediately after the program is started, the job finding rate is reduced. After some months the effect becomes smaller, implying that the outflow is somewhat higher for program participants. In the medium-long run the difference in the probability of having found work is close to zero. All three methods thus yield similar policy conclusions. The program has a negative effect in the short run and at best a zero effect in the medium long-run. Vikström (2016) finds a similar pattern in his dynamic evaluation of a Swedish work practice program. Note that we cannot estimate impacts at longer durations, such that we are not able to exclude the possibility of positive long-run effects. However, for job search assistance programs this is typically not expected (e.g. Card et al. (2010)).

Even though the implications are the same, the magnitude of the estimated impact different. The quasi-experimental estimate is largest in size, largely because the intention-to-treat effect is inflated by a small treatment share in the population. This estimate is also less precise, and the matching and timing-of-events estimates fall within its confidence interval at most durations (see panel (a) of Figure 22, in the Appendix).[23] Furthermore, the results in subsection 7.4 suggest that if the common trend assumption is violated, it most likely leads to a (small) downward bias in the quasi-experimental estimate.

9 Conclusion

Several methods are available when evaluating activation programs for unemployed job seekers. In this paper we compare estimates from three different methods. First,

[23]Since the estimators are based on different models and are not independent, the confidence intervals cannot convincingly show whether the differences between methods are significant. To test for significant differences, we bootstrap the difference between the quasi-experimental and matching estimator. The difference between the estimators, including a 95% confidence interval, is presented in Figure 24. We find that the two methods are only different at very short durations and after 11 months since the program start.

matching estimators rely on a large set of individual covariates to justify the conditional independence assumption. Second, the timing-of-events model allows for unobserved heterogeneity at the expense of a making functional form assumptions on the hazard rate specifications. Third, we exploit exogenous variation in program participation, caused by budgetary problems of the UI administration. The resulting discontinuity in program participation acts as a natural experiment.

The three resulting estimates are not identical, but reveal similar conclusions on the effectiveness of the program. All three methods suggest a significantly negative effect of program participation on outflow to employment over the first few months. This is in line with the well-documented lock-in effect. The magnitude of the negative effect differs somewhat. While the quasi-experimental estimates suggest reductions in outflow probabilities of up to 40%-points, the matching and timing-of-events estimates are somewhat smaller (5-15%-points).

At longer durations, the quasi-experimental estimates suggest an (imprecise) zero effect on employment. Both the matching and timing-of-events estimates converge towards zero at longer durations, but remain significantly negative. The similarity of the results shows that, conditional on the wide set of observed characteristics, there is no strong selection in program participation. This lack of selection is confirmed by the absence of estimated unobserved heterogeneity in the timing-of-events model. In any case, such selection is not strong enough to cause large differences in the findings. The broad conclusion drawn from each method is that the programs are not effective in increasing outflow.

Our findings are based on a specific type of individuals (those that were most likely to participate in the program), which we selected to increase homogeneity of the sample as well as maximize the treated share before the policy discontinuity. Both factors improve the performance of all our estimators. This group has relatively favorable characteristics though (native males with above average wages and below average unemployment histories), which may affect our empirical results.

Our findings can, therefore, not be generalized to the full population. Since this paper focuses on the comparison of methods rather than estimating the average effect of the program on the full population, we opted for restricting our sample in this manner.

In the meta-analysis performed by Kluve (2010), no relation is found between the methodology and the likelihood of positive or negative effects. Our results are in line with this finding, though our comparison is across methods applied to the same data, setting and program, rather than across different studies. We conclude that in the case of activation programs, a large set of observed characteristics may be sufficient to correct for most selectivity in participation.

References

Abadie, A. and Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics*, 29(1):1–11.

Abbring, J. H. and Heckman, J. J. (2007). Econometric evaluation of social programs, part iii: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6B, pages 5145–5303. Elsevier.

Abbring, J. H. and Van den Berg, G. J. (2003). The nonparametric identification of treatment effects in duration models. *Econometrica*, 71(5):1491–1517.

Abbring, J. H. and Van den Berg, G. J. (2005). Social experiments and instrumental variables with duration outcomes. Discussion Paper 05-047/3, Tinbergen Institute.

Abbring, J. H., Van den Berg, G. J., and Van Ours, J. C. (2005). The effect of unemployment insurance sanctions on the transition rate from unemployment to employment. *Economic Journal*, 115(505):602–630.

Behaghel, L., Crépon, B., and Gurgand, M. (2014). Private and public provision of counseling to job seekers: Evidence from a large controlled experiment. *American Economic Journal: Applied Economics*, 6(4):142–74.

Biewen, M., Fitzenberger, B., Osikominu, A., and Paul, M. (2014). The effectiveness of public-sponsored training revisited: The importance of data and methodological choices. *Journal of Labor Economics*, 32(4):837–897.

Brodaty, T., Crépon, B., and Fougère, D. (2002). Do long-term unemployed workers benefit from active labor market programs? Evidence from France, 1986-1998. Technical report, mimeo.

Busso, M., DiNardo, J., and McCrary, J. (2009). New evidence on the finite sample properties of propensity score matching and reweighting estimators. Discussion Paper 3998, IZA Bonn.

Card, D., Ibarrarn, P., Regalia, F., Rosas-Shady, D., and Soares, Y. (2011). The labor market impacts of youth training in the Dominican Republic. *Journal of Labor Economics*, 29(2):267–300.

Card, D., Kluve, J., and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *Economic Journal*, 120(548):F452–F477.

Card, D. and Sullivan, D. G. (1988). Measuring the effect of subsidized training programs on movements in and out of employment. *Econometrica*, 56(3):497–530.

Cockx, B. and Dejemeppe, M. (2012). Monitoring job search effort: An evaluation based on a regression discontinuity design. *Labour Economics*, 19(5):729–737.

Crépon, B., Ferracci, M., and Fougère, D. (2012). Training the unemployed in France: How does it affect unemployment duration and recurrence? *Annales d'Economie et de Statistique*, 107-108:175–200.

De Groot, N. and Van der Klaauw, B. (2014). The effects of reducing the entitlement period to unemployment insurance benefits. Discussion Paper 8336, IZA Bonn.

Dehejia, R. H. and Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94(448):1053–1062.

Dolton, P. and O'Neill, D. (2002). The long-run effects of unemployment monitoring and work-search programs: Experimental evidence from the United Kingdom. *Journal of Labor Economics*, 20(2):381–403.

Gerfin, M. and Lechner, M. (2002). A microeconometric evaluation of the active labour market policy in Switzerland. *Economic Journal*, 112(482):854–893.

Graversen, B. K. and Van Ours, J. (2008). How to help unemployed find jobs quickly: Experimental evidence from a mandatory activation program. *Journal of Public Economics*, 92(1011):2020–2035.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, 64(4):605–654.

Heckman, J. J., Lalonde, R. J., and Smith, J. A. (1999). The economics and econometrics of active labor market programs. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3A, chapter 31, pages 1865–2097. Elsevier.

Imbens, G. (2014). Matching methods in practice: Three examples. Working Paper 19959, NBER.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Kastoryano, S. and Van der Klaauw, B. (2011). Dynamic evaluation of job search assistance. Discussion Paper 5424, IZA Bonn.

Kluve, J. (2010). The effectiveness of European active labor market programs. *Labour Economics*, 17(6):904–918.

Lalive, R., van Ours, J. C., and Zweimüller, J. (2005). The effect of benefit sanctions on the duration of unemployment. *Journal of the European Economic Association*, 3(6):1386–1417.

Lalive, R., Van Ours, J. C., and Zweimüller, J. (2008). The impact of active labour market programmes on the duration of unemployment in Switzerland. *Economic Journal*, 118(525):235–257.

LaLonde, R. J. (1986). Evaluating the econometric evaluations of training programs with experimental data. *American Economic Review*, 76(4):604–620.

Lechner, M., Miquel, R., and Wunsch, C. (2011). Long-run effects of public sector sponsored training in West Germany. *Journal of the European Economic Association*, 9(4):742–784.

Lechner, M. and Wunsch, C. (2009). Are training programs more effective when unemployment is high? *Journal of Labor Economics*, 27(4):653–692.

Lechner, M. and Wunsch, C. (2013). Sensitivity of matching-based program evaluations to the ability to control variables. *Labour Economics*, 21:111–121.

Mueser, P. R., Troske, K. R., and Gorislavsky, A. (2007). Using state administrative data to measure program performance. *Review of Economics and Statistics*, 89(4):761–783.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Sianesi, B. (2004). An evaluation of the Swedish system of active labor market programs in the 1990s. *Review of Economics and Statistics*, 86(1):133–155.

Smith, J. A. and Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125(12):305–353.

Van den Berg, G. J., Bozio, A., and Dias, M. C. (2014). Policy discontinuity and duration outcomes. Discussion Paper 8450, IZA Bonn.

Van den Berg, G. J. and Van der Klaauw, B. (2006). Counseling and monitoring of unemployed workers: Theory and evidence from a controlled social experiment. *International Economic Review*, 47(3):895–936.

Van den Berg, G. J., Van der Klaauw, B., and Van Ours, J. C. (2004). Punitive sanctions and the transition rate from welfare to work. *Journal of Labor Economics*, 22(1):211–241.

Van der Klaauw, B. and Van Ours, J. C. (2013). Carrot and stick: How re-employment bonuses and benefit sanctions affect exit rates from welfare. *Journal of Applied Econometrics*, 28(2):275–296.

Vikström, J. (2016). Dynamic treatment assignment and evaluation of active labor market policies. Technical report, IFAU Uppsala.

A Timing-of-events model

The timing-of-events model contains two hazard rates, which have a mixed proportional specification. The job finding rate is given by:

$$\theta_e(t|x,\tau_0,s,v_e) = \phi_e(t)\psi_e(\tau_0+t)\exp\Big[x\beta_e + \delta_{t-s}I(t>s)\Big]v_e \tag{12}$$

and the entry rate into the program by:

$$\theta_p(s|x,\tau_0,v_p) \;\; = \;\; \phi_p(s)\psi_p(\tau_0+s)\exp(x\beta_p)v_p \tag{13}$$

We specify the duration dependence patterns $\phi_e(t)$ and $\phi_p(s)$ as piecewise constant, so $\phi_j(t) = \exp\Big(\sum_{m=1}^{M}\pi_{jm}I_m(t)\Big)$ for $j = e,p)$, where $I_m(t)$ describes duration intervals with thresholds after $3,6,9,12,18,24$ months. The calendar time indicators $\psi_e(\tau_0+t)$ and $\psi_p(\tau_0+t)$ contain dummy variables for each quarter.

The unobserved heterogeneity (v_e, v_p) is modeled using a discrete mass-point specification, so

$$p_k = \Pr(v_e = v_{ek}, v_p = v_{pk})$$

with unrestrictive mass points $v_{e1},\ldots v_{eK}$ and $v_{p1},\ldots,v_{pK}$ under the restriction $p_1 + \cdots + p_K = 1$. We try different values of K. In practice, for higher values of K, the locations of some mass points converge.

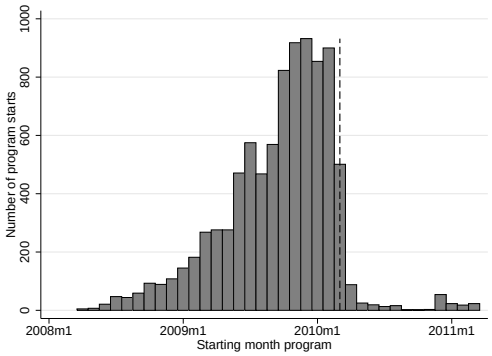
The loglikelihood function takes exogenous right censoring of durations into account. The loglikelihood function is specified as

$$\begin{aligned}\log\mathcal{L} \;\; = \sum_i \log \;\; \Big[\sum_{k=1}^{K} p_k \cdot \theta_e(t_i|x_i,\tau_{0i},s_i,v_{ek})^{c_{ei}} \cdot \exp\Big(-\int_0^{t_i}\theta_e(u|x_i,\tau_{0i},s_i,v_{ek})du\Big)\cdot \\ \theta_p(s_i|x_i,\tau_{0i},v_{pk})^{c_{pi}} \cdot \exp\Big(-\int_0^{s_i} h_a(u|x_i,\tau_{0i},v_{pk})du\Big)\Big]\end{aligned}$$

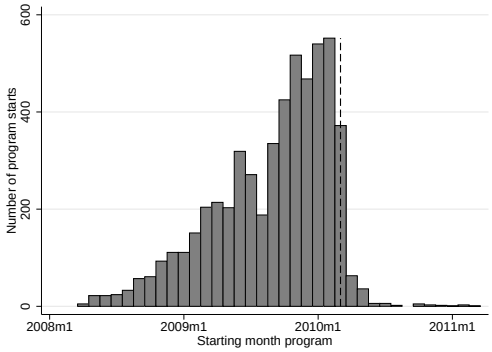
where t_i is the observed unemployment duration, s_i the observed duration before entering the program, c_{ei} denotes if the individuals finds work and c_{pi} describes if entry in the program is observed.

B Additional empirical material

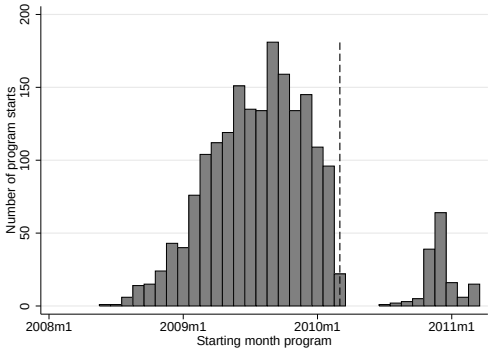
Figure 17: Starting dates of externally provided programs



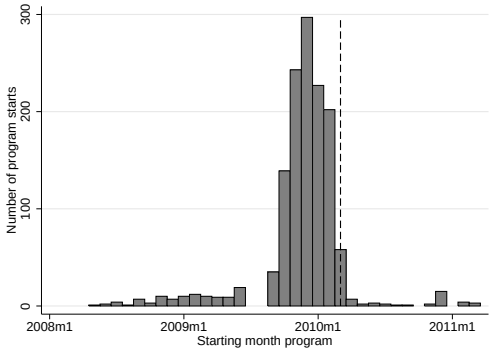
(a) IRO (Individual reintegration agreement)



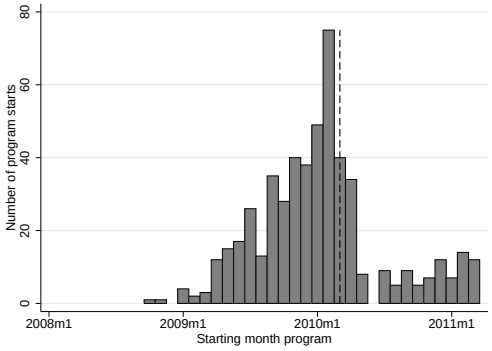
(b) Short training



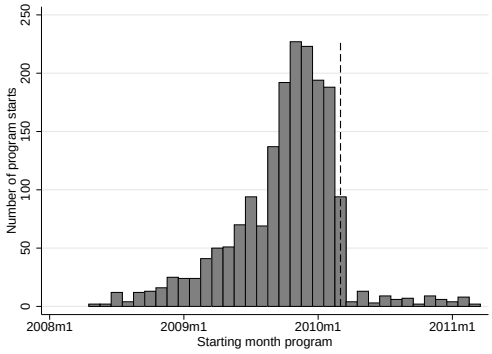
(c) Jobhunting



(d) Standard programs

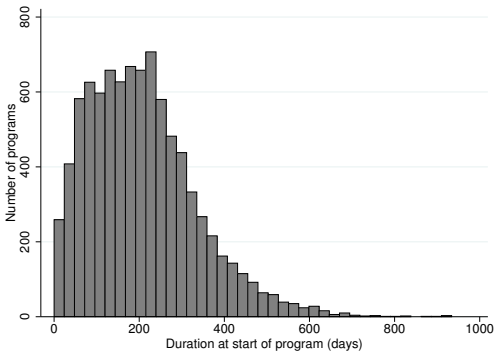


(e) Learn-work jobs

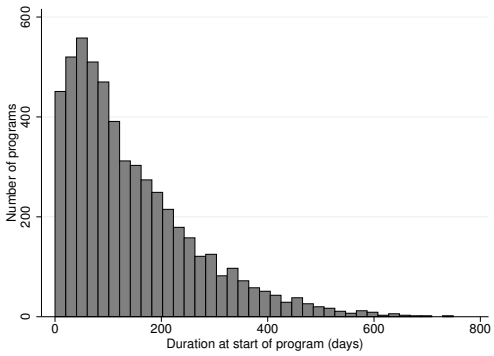


(f) Schooling

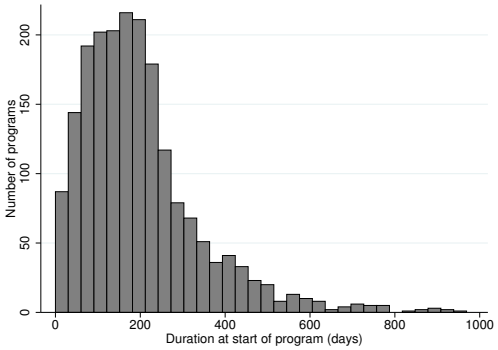
Figure 18: Timing of externally provided programs



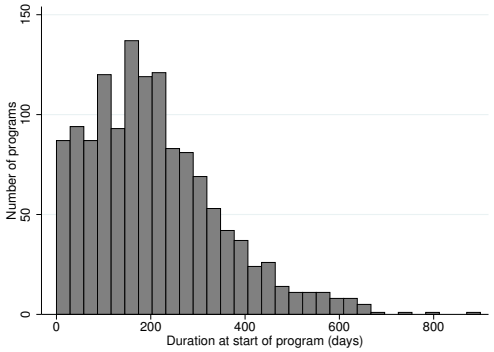
(a) IRO (Individual Reintegration Agreement)



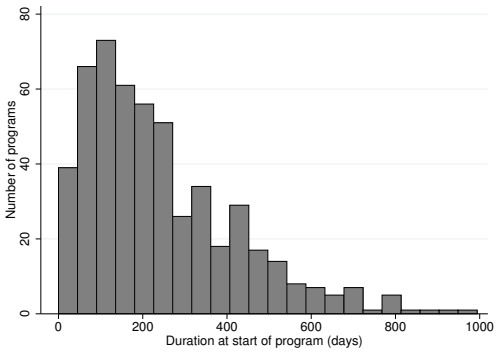
(b) Short training



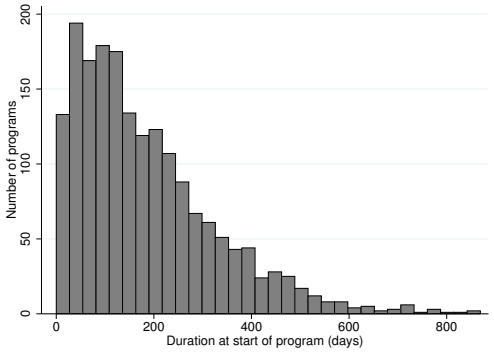
(c) Jobhunting



(d) Regular programs

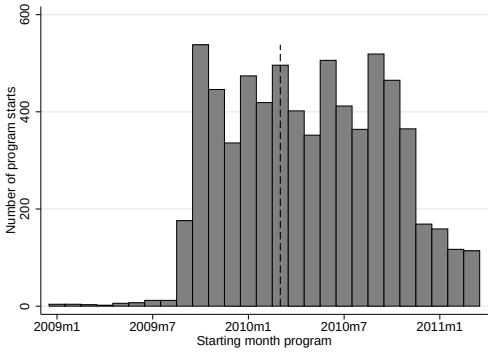


(e) Learn-work jobs

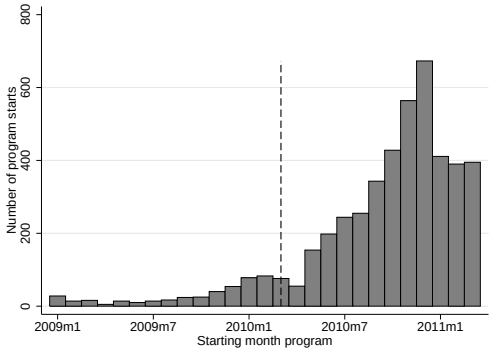


(f) Schooling

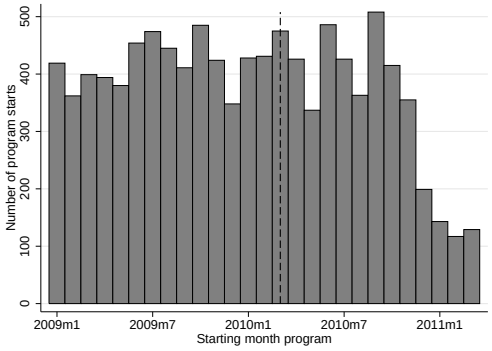
Figure 19: Starting dates of internally provided programs



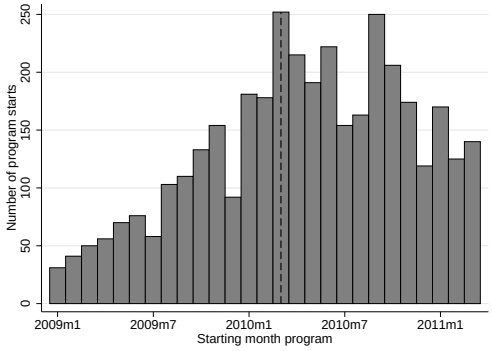
(a) CTC



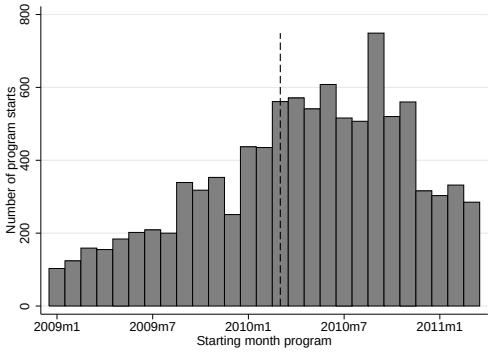
(b) Vacancy referral



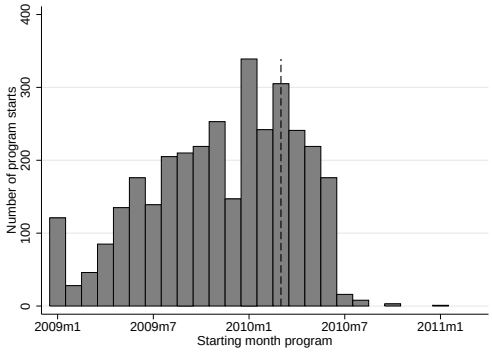
(c) Tests (different types)



(d) Employment on trial basis

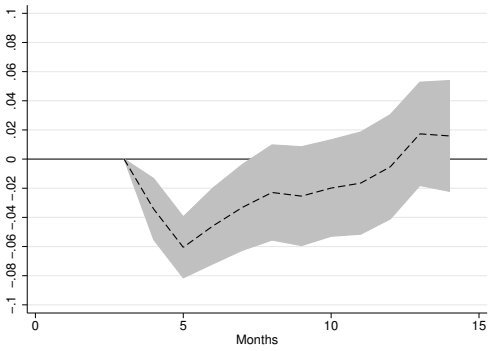


(e) Workshop

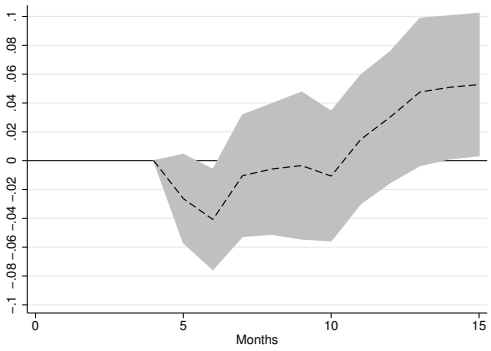


(f) Entrepeneurship support program

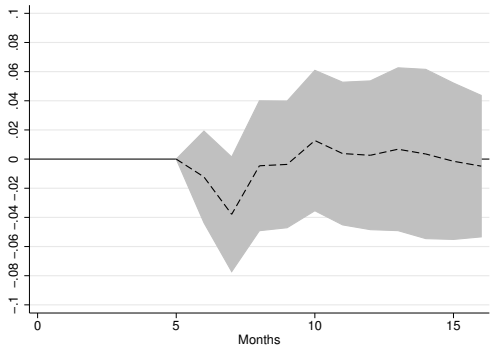
Figure 20: Experimental estimates ($\mu^2(t)$) with 95% confidence intervals



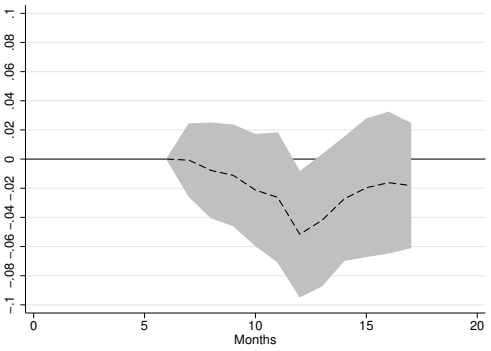
(a) Jan 2010-Oct 2009



(b) Dec 2009-Sep 2009

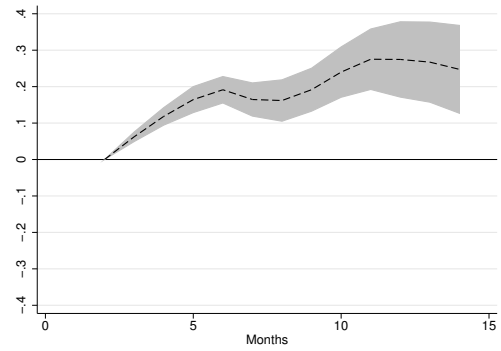


(c) Nov 2009-Aug 2009

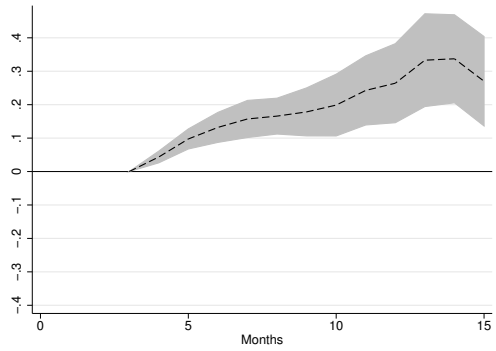


(d) Oct 2009-Jul 2009

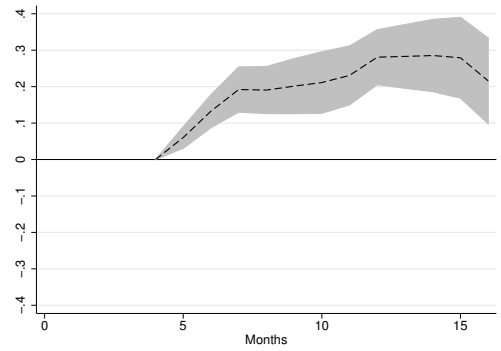
Figure 21: Double difference in treatment share with 95% confidence intervals



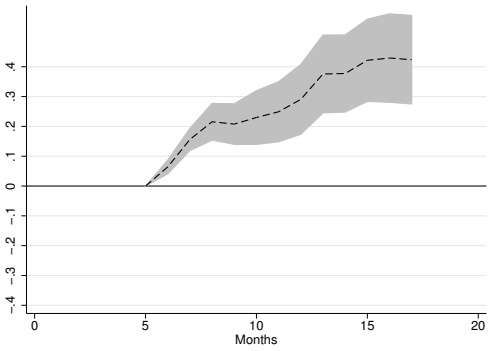
(a) Jan2010-Oct2009



(b) Dec2009-Sep2009

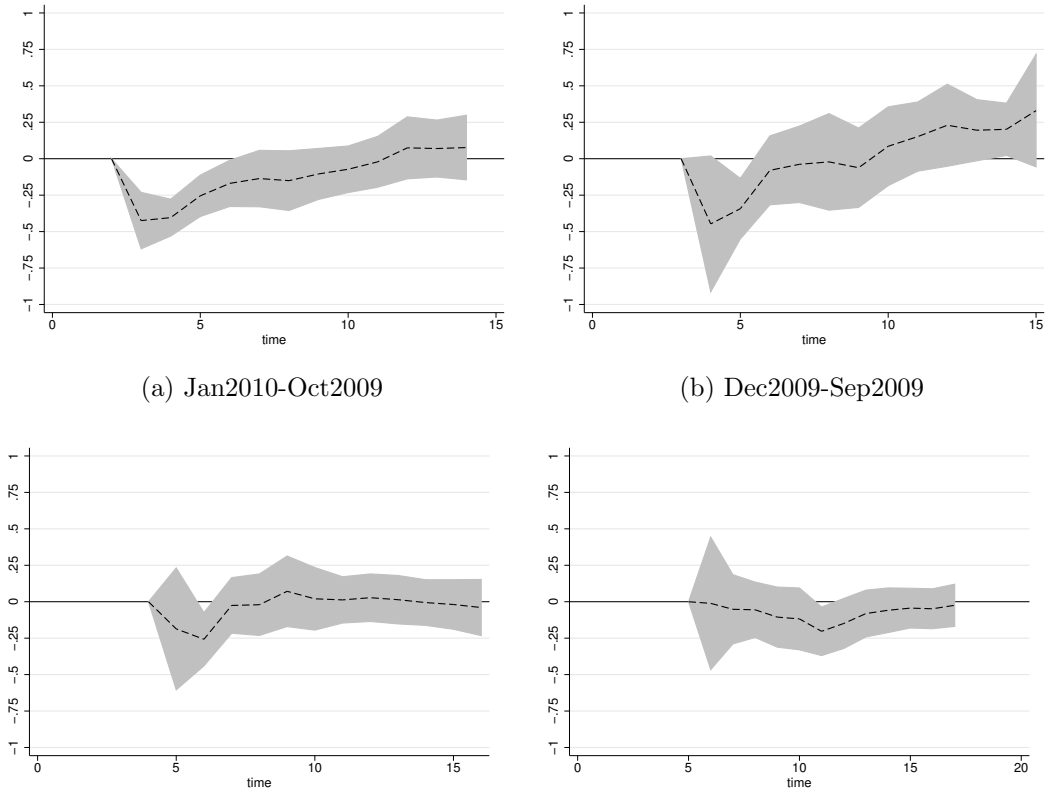


(c) Nov2009-Aug2009



(d) Oct2009-Jul2009

Figure 22: Average treatment effect with 95% confidence intervals



(a) Jan2010-Oct2009

(b) Dec2009-Sep2009

(c) Nov2009-Aug2009

(d) Oct2009-Jul2009

Figure 23: Estimated propensity scores for treated and non-treated individuals

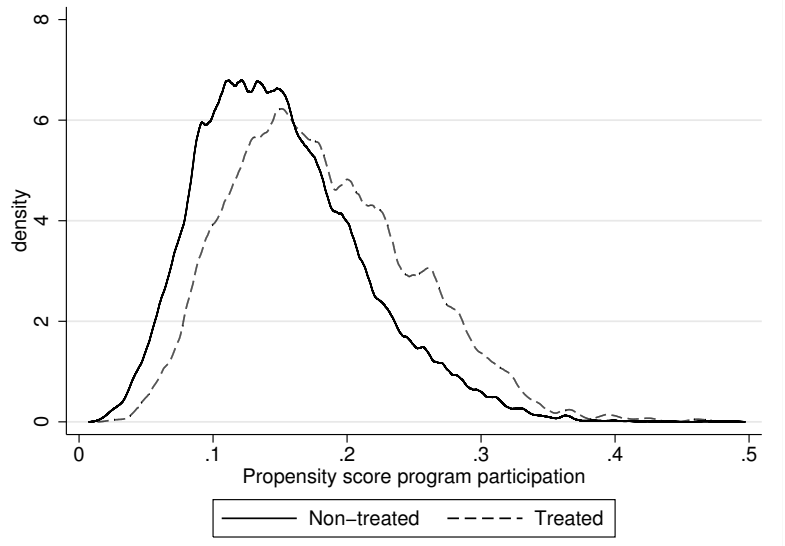


Figure 24: Difference between experimental and matching estimator (with 95% confidence interval)

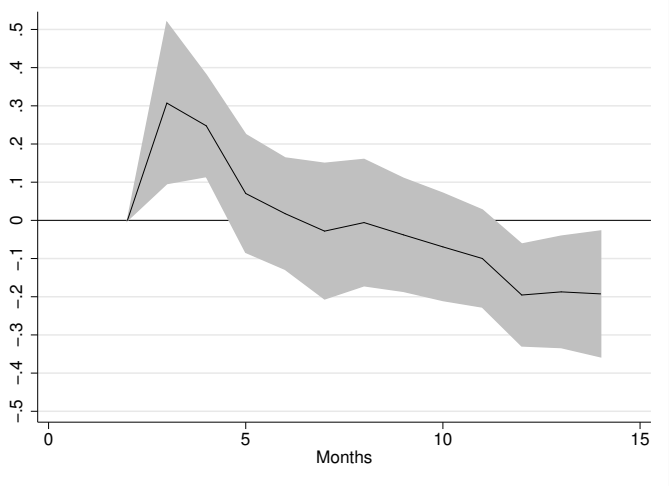


Table 4: Estimates Timing-of-Events model

	Full sample				Discontinuity sample				Pre-disc. sample			
	UI exit rate		Program rate		UI exit rate		Program rate		UI exit rate		Program rate	
	Coef.	st.er.	Coef.	st.er.	Coef.	st.er.	Coef.	st.er.	Coef.	st.er.	Coef.	st.er.
Program effect:												
Months 0-3	0.723	0.021			0.765	0.043			0.606	0.032		
Months 4-6	0.873	0.022			0.796	0.047			0.738	0.042		
Months 7-	1.080	0.016			1.046	0.036			0.853	0.056		
Individual characteristics:												
Unemployed hours	1.015	0.001	1.005	0.002	1.017	0.002	1.009	0.004	1.004	0.002	0.991	0.002
Middle educated	1.092	0.010	1.200	0.018	1.078	0.021	1.230	0.043	1.058	0.017	1.112	0.020
High educated	1.114	0.012	1.007	0.023	1.041	0.027	0.916	0.059	1.092	0.020	0.960	0.025
Income (cat. 4)	1.102	0.011	0.965	0.019	1.139	0.024	0.933	0.045	1.032	0.018	0.903	0.021
Income (cat. 5)	1.030	0.012	0.803	0.020	1.079	0.025	0.726	0.050	1.041	0.019	0.791	0.022
Age 45-50	0.924	0.010	1.093	0.019	0.902	0.021	1.048	0.048	0.906	0.016	1.110	0.022
Age 50-55	0.787	0.012	1.169	0.022	0.784	0.023	1.118	0.053	0.825	0.019	1.201	0.024
Age 55-60	0.474	0.015	0.846	0.025	0.492	0.027	0.802	0.062	0.549	0.023	0.901	0.028
UI eligibility	0.999	0.000	1.000	0.000	0.999	0.000	1.000	0.000	0.999	0.000	1.000	0.000
Region 2	0.931	0.014	0.793	0.024	0.908	0.029	0.858	0.053	0.875	0.024	0.710	0.027
Region 3	0.987	0.014	0.702	0.024	0.954	0.029	0.657	0.056	0.866	0.024	0.631	0.027
Region 4	0.903	0.016	0.759	0.027	0.829	0.034	0.676	0.066	0.885	0.026	0.714	0.030
Region 5	0.953	0.015	0.612	0.025	0.909	0.030	0.527	0.063	0.925	0.024	0.544	0.029
Region 6	0.929	0.015	0.672	0.026	0.903	0.031	0.576	0.063	0.910	0.026	0.592	0.030
Region 7	1.026	0.030	0.280	0.082	1.051	0.064	0.516	0.155	1.243	0.046	0.303	0.086
UI history (cat. 2)	1.262	0.011	1.015	0.022	1.309	0.024	0.820	0.055	1.182	0.019	0.977	0.025
UI history (cat. 3)	1.081	0.013	0.997	0.024	1.173	0.028	0.820	0.063	0.956	0.022	0.944	0.027
UI history (cat. 4)	0.827	0.019	0.956	0.029	0.860	0.048	0.982	0.089	0.746	0.029	0.870	0.032
UI history (cat. 5)	0.804	0.031	0.896	0.045	0.818	0.088	1.075	0.138	0.727	0.043	0.748	0.050
Sickness or disability history (short)	0.631	0.034	1.084	0.050	0.567	0.081	1.447	0.118	0.574	0.062	1.062	0.056
Sickness or disability history (long)	0.403	0.033	0.918	0.041	0.355	0.085	1.028	0.116	0.337	0.060	0.847	0.047
Sickness or disability during UI	0.736	0.030	0.900	0.045	0.811	0.064	1.202	0.105	0.698	0.050	0.691	0.055
Cohabiting	1.212	0.015	0.973	0.026	1.205	0.033	0.993	0.065	1.173	0.024	0.893	0.029
Married	1.387	0.012	0.988	0.020	1.423	0.027	1.041	0.052	1.312	0.019	0.956	0.022
Single parent	1.053	0.035	0.946	0.060	1.003	0.087	0.848	0.177	0.981	0.055	0.862	0.067
Divorced	1.070	0.017	0.960	0.028	1.119	0.037	0.920	0.072	0.989	0.028	0.927	0.031
Unknown marital status	1.013	0.032	1.095	0.064	0.917	0.075	1.223	0.144	0.863	0.059	0.894	0.076
Duration dependence:												
Months 1-3	1		1		1		1		1		1	
Months 4-6	1.319	0.011	1.371	0.018	1.140	0.029	1.031	0.046	1.443	0.017	1.475	0.019
Months 7-9	1.053	0.013	1.788	0.020	0.867	0.045	0.285	0.318	1.298	0.024	2.108	0.023
Months 10-12	0.810	0.016	1.881	0.025	0.979	0.064	0.100	0.784	0.831	0.039	1.874	0.033
Months 13-18	0.659	0.017	1.427	0.032	0.966	0.091			0.766	0.051	1.629	0.047
Months 19-24	0.534	0.022	0.990	0.079	1.186	0.121			0.452	0.141		
Months 25-	0.514	0.030	0.131	1.094	1.515	0.177						
Calendar time effects												
Quarter 1	1		1						1		1	
Quarter 2	0.866	0.040	0.426	0.067					0.793	0.039	0.194	0.064
Quarter 3	0.506	0.040	0.561	0.058					0.439	0.040	0.301	0.049
Quarter 4	0.581	0.037	0.573	0.054					0.529	0.037	0.270	0.046
Quarter 5	0.509	0.036	0.633	0.053					0.456	0.036	0.341	0.043
Quarter 6	0.467	0.036	0.650	0.052					0.377	0.035	0.443	0.040
Quarter 7	0.324	0.037	0.925	0.051	1		1					
Quarter 8	0.597	0.035	0.711	0.052	3.703	0.045	0.867	0.039				
Quarter 9	0.635	0.035	0.030	0.083	4.713	0.054	0.065	0.130				
Quarter 10	0.564	0.035	0.003	0.194	3.470	0.067	0.038	0.436				
Quarter 11	0.477	0.037	0.008	0.134	2.011	0.086	0.091	0.818				
Quarter 12	0.666	0.037	0.006	0.169	2.580	0.104	0.428	0.803				
Quarter 13	0.643	0.039	0.002	0.254	1.858	0.113	0.081	0.823				
Quarter 14	0.559	0.042			1.335	0.136						
Quarter 15	0.412	0.047			0.838	0.154						
Quarter 16	0.491	0.062			0.735	0.215						
Unobserved heterogeneity												
v_1	0.015		0.022		0.000		0.001		0.003		0.000	
v_2	0.003		0.002		0.997		0.368		0.007		0.009	
p_1	0.002				1.000				0.053			
p_2	0.998				0.000				0.947			
Observations	116,652				23,502				83,773			

All reported values are hazard rates.