Comparing Econometric Methods to Empirically Evaluate Job Search Assistance

Paul Muller* Bas van der Klaauw** Arjan Heyma***

February 22, 2019

Abstract

We test whether different identification strategies 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 (1) exploiting the policy discontinuity as a quasi-experiment, (2) using dynamic matching assuming conditional independence, and (3) applying the timing-of-events model. All three strategies use the same data to consider the same program in the same setting, and show that the program reduces job finding directly after enrollment. However, the magnitude of the estimated drop in job finding differs between the three estimation methods. In the longer run, all three methods show a zero effect on employment.

Keywords: policy evaluation, job search assistance, unemployment duration

JEL-code: J64, C52, C41

*VU University Amsterdam

**VU University Amsterdam and Tinbergen Institute

***SEO Economic Research

VU University Amsterdam, Department of Economics, De Boelelaan 1105, 1081 HV Amsterdam. Email: p.muller@vu.nl, b.vander.klaauw@vu.nl, a.heyma@seo.nl.

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 the EALE (2013), ESPE (2014), VU University Amsterdam, University of Edinburgh, BeNA Berlin, University of Vienna, University of Copenhagen, University of Gothenburg, DIW Berlin, IZA Bonn and Collegio Carlo Alberto Turin. Van der Klaauw acknowledges financial support from a Vici-grant from the Dutch Science Foundation (NWO).

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 of the programs on job finding. In addition, we estimate the same effects using non-experimental estimators (matching and timing-of-events) and compare the results of the three identification strategies.

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.[1] Alternatively, some studies exploit institutional features or policy changes that generate random variation in program participation.[2] Less than 10% of the studies use an experimental design.[3] LaLonde (1986) shows that non-experimental estimators produce

[1]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) look at long-run effects of training programs in Germany.

[2]For example, 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 a 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, Van den Berg and Van der Klaauw (2006) analyze a randomized experiment on counseling and monitoring in the Netherlands, Graversen and Van Ours (2008) evaluate an experiment on an intensive activation program in Denmark, Card et al. (2011) estimate the impact of a training program in the Dominican Republic and Behaghel et al. (2014) perform a large

results that do not concur with those from experimental evidence. While Dehejia and Wahba (1999) find that alternative matching estimators produce results closer to the experimental evidence, their findings are disputed. Smith and Todd (2005) show that these findings are not robust to different specifications and the use of different samples and different sets of covariates. Experimental and non-experimental estimates can only be expected to produce similar results if three requirements are fulfilled (Heckman et al. (1997)). First, the data source should be the same for the treatment and control group (in particular the outcome variable should be measured in the same way). Second, treated and control individuals should be 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 contribute to this literature by performing a similar comparison of identification strategies, using administrative data. Our 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 detailed information on the timing of privately provided job search assistance programs. 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 use a recently proposed dynamic matching estimator (Vikström (2017)) as well as the timing-of-events model (Ab-

controlled experiment, randomizing job seekers across publicly and privately provided counseling programs. Also using a randomized experiment, Cottier et al. (2015) find negative long-run effects of participating in programs of a large private job search assistance provider in Switzerland.

bring and Van den Berg (2003)).[4]

We exploit the discontinuity in the provision of job search assistance programs by private providers, to estimate how program participation affects job finding.[5] 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, employment probabilities are reduced significantly for several months. After half a year these increase, up to a zero difference in employment 12 to 18 months after enrollment in the program.

We compare these results to those from a dynamic matching estimator. Where the quasi-experimental estimator compares job finding between different cohorts, the matching estimator compares treated and (observationally) similar non-treated individuals within the same cohorts. The results show a significantly negative effect of program participation directly after entering the program. Even though the negative effect decreases in magnitude over time, it remains significantly negative up to 12 months after the programs starts. Next, we estimate a timing-of-events model, 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 again that the program reduces the job

[4]Note that the meta-analyses by Card et al. (2010) and Kluve (2010) also pose the question of whether a relationship exists between the identification strategy and the empirical results. While they find no evidence for such a relationship, their comparison is across studies and thus cannot compare different estimators applied to the same setting. Some recent contributions to the job search assistance literature have focused on comparing different estimators. See for example Mueser et al. (2007), Lalive et al. (2008), Biewen et al. (2014) and Kastoryano and Van der Klaauw (2011). Comparisons of (quasi-)experimental and non-experimental estimators are still rare though.

[5]We also refer to the program discontinuity as a “quasi-experiment”.

finding rates 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 year, and a zero effect afterwards.

Thus, we find that in our setting, quasi-experimental and non-experimental identification strategies lead to fairly similar conclusions: the program reduces job finding in the short and medium run while there is no impact afterwards. The magnitude of the negative effect varies across the different estimators, but the main policy recommendations are the same.[6] The dynamic matching estimator and the timing-of-events estimator can use a more extensive sample than the quasi-experimental approach, but results are very robust against extending the data set. The finding that treatment effects estimated using the timing-of-events model are closer to zero than when estimated using dynamic matching suggests that there may be some selection on unobservables and in particular more disadvantaged unemployed workers are more likely to participate in the program. The quasi-experimental treatment effects are estimated much less precise than when using both other approaches, mainly because intention-to-treat effects are inflated with small shares of program participants.[7] The finding that assistance provided through the private market is ineffective, is in line with earlier evidence (Behaghel et al. (2014)).

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 job finding rates. Furthermore, results may depend on the Dutch setting and the contents of the program that we evaluate. It contains a mix of caseworker meetings, job referrals and goal setting in the job search process. These elements are common in job search assistance programs in other countries, but

[6]The negative impact on job finding in the short run is in line with the well-documented lock-in effect. In our setting, rather than being full-time occupied with the program, the lower job finding might follow from 'waiting' to see what the program has to offer, once participation has been decided.

[7]In absolute terms the monthly number of individuals entering the program is large, but this remains a small fraction of all eligible UI benefits recipients in the Netherlands.

likely with a different intensity. Other program characteristics are similar to other countries. The job search assistance is offered in addition to the "basic" assistance of the UI administration (mostly irregular meetings with caseworkers), which is the case in many countries. Also the set-up in which caseworkers have substantial discretion when assigning job seekers to programs is a feature common to many UI administrations. Finally, the Dutch 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 describe the institutional setting and the budgetary problems which led to the policy discontinuity in Section 2. An overview of the data is provided in Section 3. In Section 4 we define our treatment effect of interest. In Section 5 we present non-experimental results from the matching and timing-of-events estimators. In Section 6 we discuss how the discontinuity allows identification of the treatment effect and present estimation results. Section 7 compares the results from the different approaches and provides a discussion. Section 8 concludes.

2 Institutional setting and the policy discontinuity

In the Netherlands, UI is organized at the nationwide level. The UI administration (UWV) pays unemployment benefits to workers that 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 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 by 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 (2019) 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. Benefits recipients are obliged to accept any suitable job offer.[8] Caseworkers are responsible for monitoring compliance to these obligations, but the intensity of meetings is low. In 2009, caseworkers had the possibility of assigning an individual to a range of programs aiming at increasing the job finding rate, if they judged that the benefits recipient required more than the usual guidance. A large diversity of programs 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. These were for-profit companies and they faced payment schemes that consisted partly of lump-sum payments and partly of performance payments. The performance payments were not very sophisticated and varied between contracts. Often an additional payment would be made if a certain share of the program participants would leave the UI benefits system within a particular time period.

Our analysis focuses solely on the programs that were externally provided by private companies. 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 the assignment of the different programs.

The lack of centralized program assignment together with an increased inflow in

[8] During the first six months of unemployment 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.

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.[9] Assistance offered internally by the UI administration continued without change. In Section 3 we show that indeed the number of new program entrants dropped to almost zero in March 2010 and remained very low afterwards.

3 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 583,193 observations (each UI spell is considered an observation, though for some individuals there are multiple spells).[10]

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 benefits payments. Right censoring occurs on January 1st, 2012, when our data were 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

[9]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).

[10]The original dataset contains 671,743 unemployment spells. We exclude 34,968 spells from individuals previously employed in the public sector and 17,454 from individuals older than 60 years. Next, we drop 25,778 spells for which important variables are missing. We also exclude 524 spells from individuals working more than 60 hours or less than 12 hours in their previous job and 9,504 spells during which the job seeker also had ongoing employment. Finally we exclude 288 spells with a negative unemployment duration and 34 spells with negative benefits eligibility.

Table 1: Descriptive statistics

	Full sample	Inflow cohort UI: April 2008	April 2009	April 2010
Unemployment duration (median, days)	147	123	161	158
Reason for exit (%):				
Work[a]	56.5	58.1	53.3	58.9
End of entitlement period	31.9	27.7	35.4	28.2
Sickness/Disability	4.6	4.9	4.7	5.1
Other	7.0	9.3	6.7	7.8
Participation external program (%):				
Any program	13.3	16.1	22.3	0.2
Job search assistance[b]	7.6	10.8	14.3	0.1
Training[c]	4.4	5.3	6.9	0.1
Subsidized employment	0.3	0.1	0.4	0.0
Other	3.1	2.8	4.6	0.0
Participation internal program (%):				
Any program	27.6	6.9[d]	31.0	31.5
Tests	8.1	1.0[d]	9.6	9.5
Subsidized employment	1.6	0.4[d]	1.7	2.1
Workshops entrepeneurship	2.6	1.0[d]	3.5	2.2
Job search assistance	7.1	0.8[d]	6.9	9.9
Other	15.2	4.8[d]	17.9	15.5
Gender (% female)	41.5	50.8	42.0	44.0
Age	37.9	38.8	37.7	38.3
Immigrant (%)	8.8	10.0	8.7	8.3
Education (%):				
Low	30.8	33.0	29.5	28.7
Middle	46.9	44.1	46.7	46.7
High	22.3	22.9	23.7	24.6
UI history last 3 year (%)	31.6	38.7	28.3	28.3
Previous income (%):				
Low	20.4	26.1	20.5	18.9
Middle	63.1	58.0	63.4	62.7
High	16.4	15.9	16.1	18.4
Unemployment size (hours)[e]	33.7	32.9	34.0	33.4
Observations	583,193	10,422	22,069	20,107

[a] Including exits due to unknown reasons. [b] Job search assistance contains 'IRO' (Individual reintegration agreement), 'Job hunting' and 'Application letter'. [c] Training contains 'Short Training' and 'Schooling'. Subsidized employment contains 'Learn-work positions'. [d] Biased downwards, because participation in internal programs was rarely recorded before 2009. [e] 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 last job.

Figure 1: Number of UI entrants per month

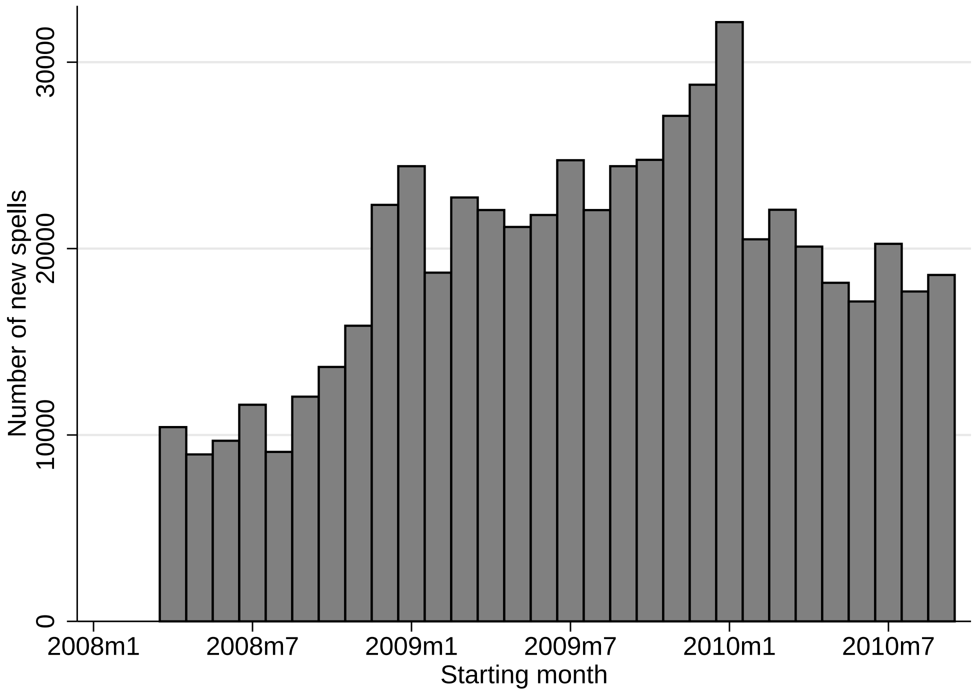
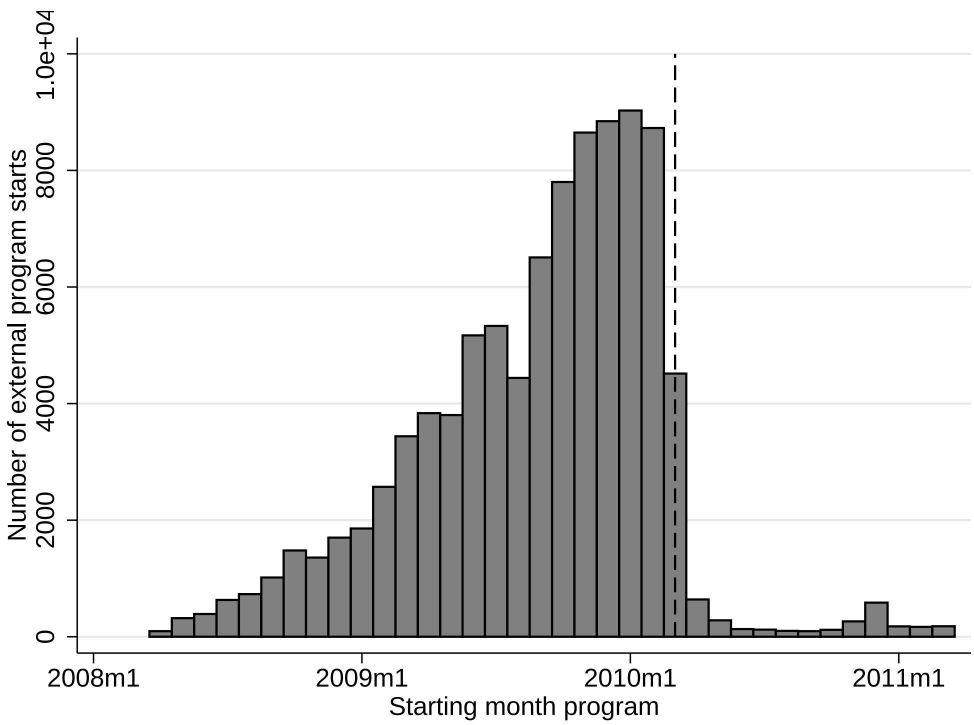


Figure 2: Number of UI benefits recipients entering an external program per month



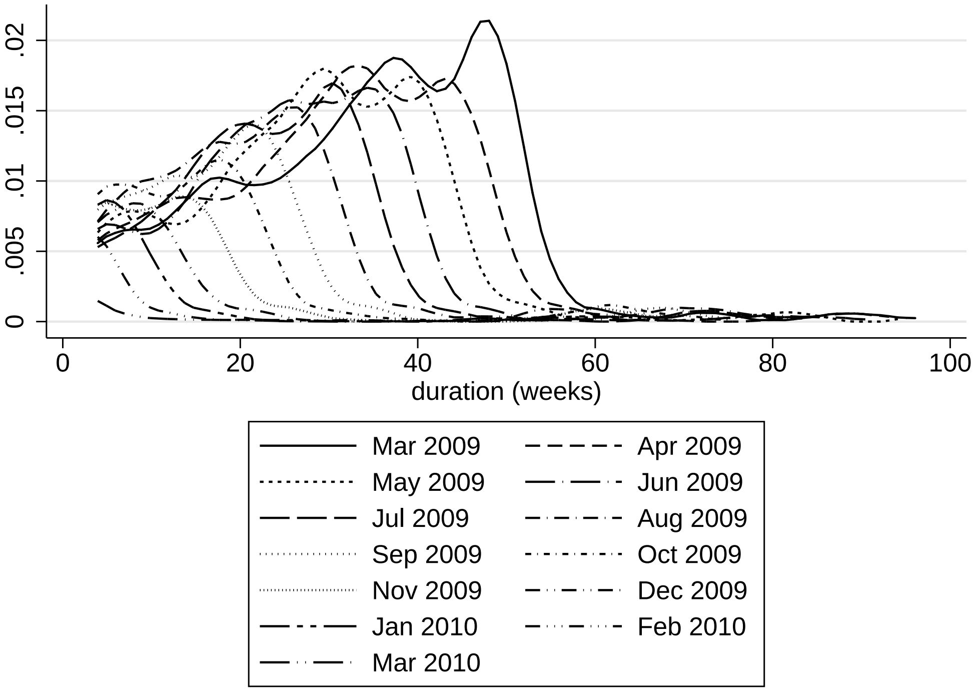
December 2008 onward. The inflow increased from about 2000 to 5000 per month and remained high until the end of 2009. From 2010 onward the inflow decreased somewhat. Table 1 presents summary statistics for the full sample, as well as for three subgroups defined by their month of inflow into unemployment. Column (1) shows that for the full sample the median duration of unemployment is 245 days (around eight months). About 70% of those exiting UI find work, which also includes spells for which the reason for outflow is unknown.[11] The end of the benefits entitlement period is reached for 15% of the spells. Almost 7% leave unemployment due to sickness or disability; the rest leave for other reasons. Exits due to reaching the end of the entitlement period and exits due 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 as our outcome of interest.

In the full sample about 19% of the benefits recipients participate in one of the externally provided programs. Over half 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 35% 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 (hours per week), 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 experienced a period of unemployment in the three years before entering UI.

[11] Often UI recipients do not give a reason for voluntarily stopping collecting benefits or these reasons are not registered by caseworkers. In most of these cases the individual starts generating income from other sources related to working.

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



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 timing 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 C1) 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 of starting an external program.[12] Each cohort reaches the policy discontinuity

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

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 (offered by the UI administration itself) remained unaffected around March 2010, limiting the scope for scaling up internal programs. The number of internal programs started in each month is shown in Figure 4. 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 C3). 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.[13]

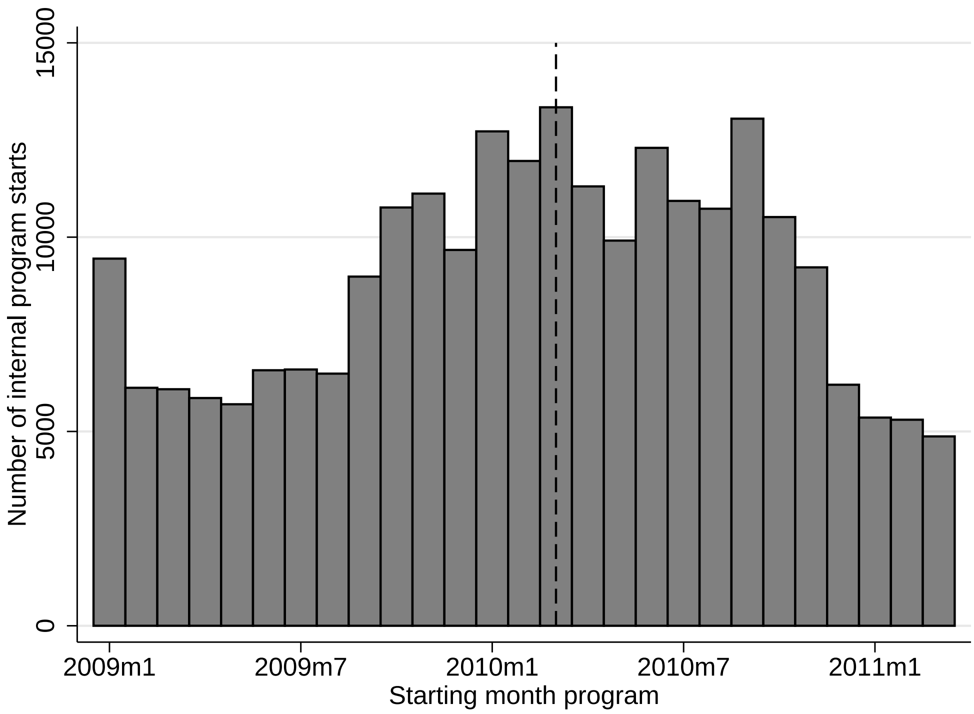
A further concern might be that even though the number of internal programs was not changed, caseworkers may have reacted to the unavailability of external programs by shifting their internal programs to these individuals that might otherwise have participated in external programs. This would imply that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

individual.

[13]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 workers such that this never happens in the Netherlands.

Figure 4: Distribution of starting dates of the internally provided programs (the dashed line indicates the policy discontinuity in March 2010)

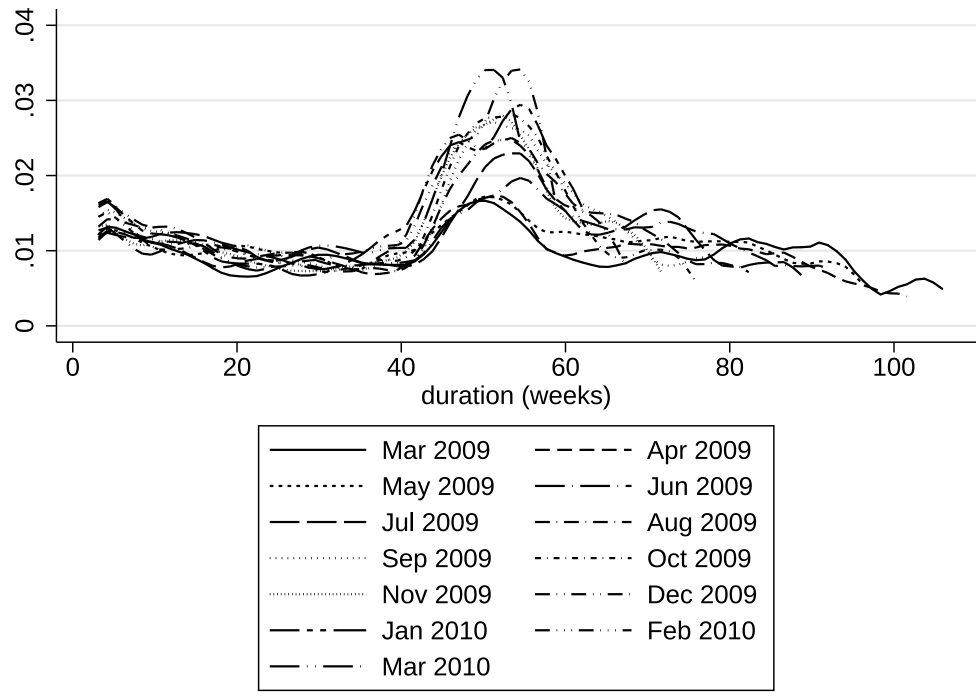


program per month in Figure 6. Mean age and weekly hours of unemployment are shown in panel (a) of Figure 6, unemployment and disability history and education level are shown in panel (b), the previous hourly wage is shown in panel (c) and the share of nine industry categories is shown in panel (d). 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.

4 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 selectivity in the participation decision, the

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

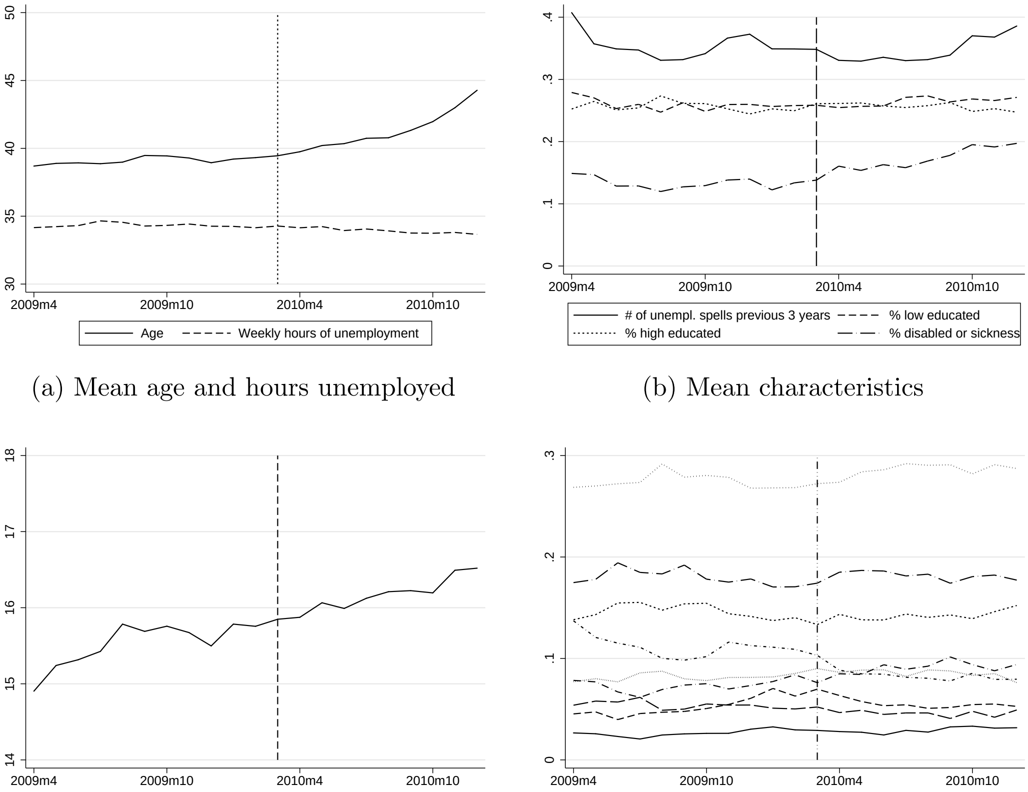


composition of program participants and non-participants is different. We focus on treatment effects for participants. These treatment effects are nonstandard because enrollment in the program is dynamic.

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 found work prior to time t, and zero if the individual is still unemployed. And define S to be a random variable denoting the duration at which program participation starts. Potential unemployment durations $T^*(s)$ depend on program participation $S = s$. Since program assignment is dynamic, even for a single type of 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:

Figure 6: Composition of the internal program participants



(a) Mean age and hours unemployed

(b) Mean characteristics

(c) Mean hourly wage (in Euro)

(d) Industry shares

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

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 $t > s$, which is necessary to define counterfactuals and thereby treatment effects. The no-anticipation assumption 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 change their job search behavior prior to s in anticipation to learning that they will enroll in the program at time s.[14] Such behavior may be unlikely for the external programs we consider in this paper. Programs are assigned by caseworkers on an individual basis. There are no strict criteria for participation and 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 a 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 the the treatment effect on the treated, that is, the effect of program participation on individuals treated at duration s. Since individuals exit from unemployment at any duration, we condition on remaining unemployed until at least duration s. This (conditional) contrast between the potential outcome under treatment at s and no treatment is defined by Van den Berg et al. (2014) as the average treatment effect on the treated survivors (ATTS(s,t)):

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

The average treatment effect on the treated survivors ATTS(t,s) provides a series of effects for different values of s and t. Policy makers often focus on all participants, which requires knowledge about the treatment assignment policy. Let $f(s)$ denote the density function of starting treatment at time s and $S_0(s)$ the survivor function in unemployment until time s for individuals who did not participate in treatment before s. We define the average treatment effect on the treated evaluated t time

[14]The no-anticipation assumption does not rule out that treatment assignment probabilities differ between individuals and that individuals are aware of this. It also allows that program enrollment is more likely in some periods than in other periods, which may be known by job seekers. Only the exact timing of the program start should be unanticipated.

periods after entering unemployment as

$$\text{ATET}(t) = \frac{\int_{s=0}^{t} f(s) S_0(s) \text{ATTS}(s,t) ds}{\int_{s=0}^{t} f(s) S_0(s) ds} \tag{2}$$

This describes the average treatment effect on those individuals who started participating in the treatment before time t. For tractability of the comparison of the different estimation approaches, we will focus on estimating the ATTS(t, s), but in addition we will provide an estimate of the ATET(t) in Section 7.[15]

4.1 Choice of samples

The different identification strategies that we apply, require sample selections that do not necessarily coincide. 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. To apply each identification strategy as they would be applied in a usual application, we construct two subsamples from the full sample (see Table 2).

A difference in results can be due to different sample selection rules. We argue that choice of sample is an essential part of an identification strategy. However, to investigate to what extent the sample choice causes differences in 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 6, where we discuss the quasi-experimental approach exploiting

[15]Note that Vikström (2017) defines an alternative aggregated effect, which is the average impact τ periods after *treatment* starts:

$$\text{ATET}(\tau) = \frac{\int_{s=0}^{s+\tau} f(s) S_0(s) \text{ATTS}(s, s+\tau) ds}{\int_{s=0}^{s+\tau} f(s) S_0(s) ds} \tag{3}$$

This provides an average impact for each elapsed duration since entering a program, while in equation (2) we define the average impact for each elapsed duration since entering unemployment. We focus on our measure, because we believe it provides a more intuitive measure of the average impact of the program. Note, however, that for our estimates this alternative approach yields very similar results.

Table 2: Identification strategies 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
Observations	583,193	112,678	428,160

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

the policy discontinuity. In addition, we apply the matching and timing-of-events estimators 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 censors all observations at the time of the discontinuity. The rationale for analyzing 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.

5 Non-experimental analysis

5.1 Matching estimator

We start the empirical analysis by applying a matching estimator. The identification strategy does not exploit the policy discontinuity, but instead compares individuals with similar characteristics differing only in treatment status. We apply a dynamic matching estimator (proposed by Vikström (2017)) to account for the dynamic setting and dynamic selection.

Matching methods construct counterfactual outcomes in the absence of randomized assignment (Imbens and Wooldridge (2009), Heckman et al. (1997)). The approach relies on two main assumptions. First, selection into treatment is on observables only:

$$Y_{0,t}^*(s), Y_{1,t}^*(s) \quad \perp S|X = x \tag{4}$$

This unconfoundedness assumption implies that after conditioning on a set of observed characteristics, assignment to treatment is independent of the potential outcomes. See Vikström (2017) for a discussion of how this assumption generalizes to a setting of dynamic treatment assignment. Our administrative data include a rich set of covariates, which is crucial for the likelihood of unconfoundedness assumption to hold. Employment histories are argued to be particularly important, because they tend to be 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 at 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 nonparticipants. For our dynamic setting we assume

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

where $f(s; x)$ 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.

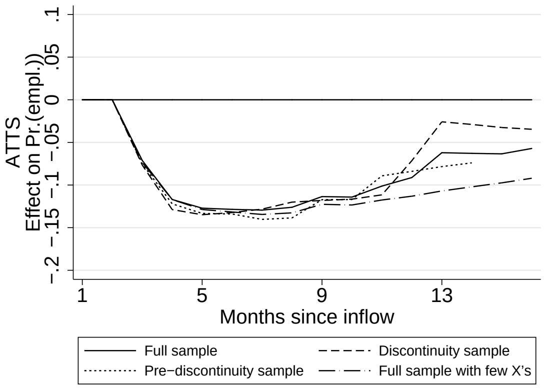
As a baseline, we consider the full sample, and as robustness checks we use two subsamples (as discussed in Subsection 4.1). We compare those that are treated at s with those that do not get treatment. Clearly, the probability of never receiving treatment depends on the length of an unemployment spell. That is, the longer someone remains unemployed, the larger the probability of starting treatment. To correct for the dynamic selection, we apply the estimator proposed by Vikström (2017).[16] The treatment group contains all individuals that start treatment at s. The control group for the impact of the treatment on outflow before duration $s + t$ contains all individuals that have not been treated and are still unemployed by duration s. As soon as an individual in the control group is assigned to a program, her spell of unemployment is censored. Since program assignment is not random, the surviving control group individuals are weighted using propensity scores, making them comparable to the treatment group in terms of observed characteristic. Given the continuously changing composition of the control group, the weighting is repeated each period. For further details on the estimator we refer to Vikström (2017). Note that we estimate the version of the estimator that, in addition to correcting for selective exits, also corrects for potentially selective right-censoring.

We aggregate all spell data into monthly intervals and estimate the impact of treatment on job finding for different values of s, that is, programs starting at different elapsed durations.[17] The estimates are presented in Figure 7 for treatment starting at $s = 3, 4, 5, 6$ months, and measured up to an unemployment duration of 16 months. We focus on these four values of s because inflow into programs is highest at these durations (see Figure C2 in the appendix). The presented estimates are

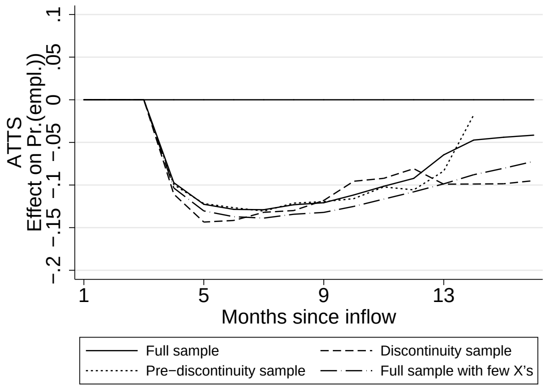
[16] An alternative approach that has been applied in the literature is to compare those treated at s with those not *yet* treated at s and interpret the difference as the effect of early treatment relative to potential later treatment (see for example Sianesi (2004)). We favor the approach of using those not getting treatment as the control group, because we believe that it has a more straightforward interpretation.

[17] To be precise, we aggregate the spell data into 30-days intervals.

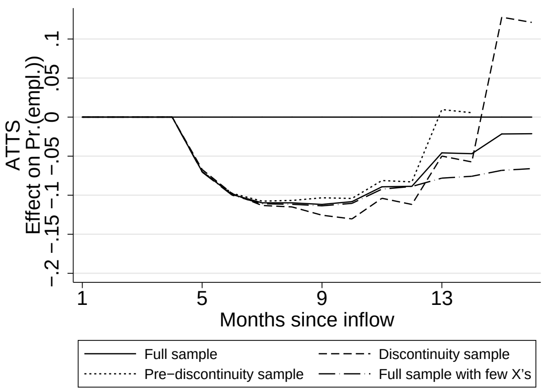
Figure 7: Dynamic matching ATTS estimates: matching estimates based on different samples



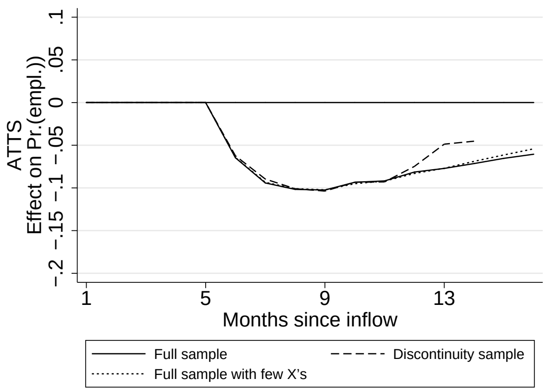
(a) s = 3



(b) s = 4



(c) s = 5



(d) s = 6

differences in survivor probabilities between the treatment group and the (weighted) control group.

The results point to a number of interesting findings. First, and most importantly, program participants are less likely to find employment. Around two to three months after the program has started, the probability of finding employment in the control group is 10 to 20%-points higher than in the treatment group. This difference remains for at least 12 months.[18] The negative impact of program participation during the first couple of months is consistent with a lock-in effect. Schooling and

[18]We present separate graphs for each value of s and each sample in Figures C4 - C7 in the appendix, in which we include confidence intervals. The negative effect is statistically significant for all estimates, and only at durations of 12 months are some estimates no longer significantly different from zero.

training programs often require substantial effort from participants which cause that unemployed workers invest less time in job search. Job search assistance programs may crowd out certain types of job search effort which are more effective for finding work (e.g. Van den Berg and Van der Klaauw (2006)).

Second, the pattern of program effects is very similar across the four panels (a)-(d) and estimated impacts for other values of s ($s = 1, \ldots, 8$) show very similar results (see Figure C4 in the appendix). The effects of the program are thus equally negative when a benefits recipient enters the program relatively fast or later during the period of unemployment.

Third, Figure 7 shows that using different samples yields very similar estimates for the impact of program participation. In the discontinuity sample selectivity in program participation should be less of an issue than in the other samples. The similarity of estimated effects can either imply that the set of covariates is sufficiently rich to deal with selective program assignment or that programs are assigned relatively random. Therefore, we repeat the matching estimator, but only using age dummies as observed covariate. This hardly affects the estimates, which provides evidence that selectivity on observed covariates is not large when assigning benefit recipients to the programs. Whether selectivity on unobserved characteristics is important when assigning the program requires comparing the findings to those from the other empirical approaches.

We conclude that the negative impact of the program on employment probabilities is (1) the same for assignment at different durations, (2) very robust across different samples and (3) even robust against excluding most observed characteristics from the matching estimator.

5.2 Timing of events model

Matching requires very modest 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 unobserved variables, but makes stronger functional-form assumptions. This model has been applied regularly in the recent literature on dynamic treatment evaluation.[19]

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. Since the model is a continuous-time model, we use daily spell data and do not have to make an assumption on which unit of time to take when discretizing the unemployment durations.

We present a concise description of the model here, while a detailed version is presented in the Appendix A. 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 a piece-wise constant function of the elapsed duration since starting the program (see Appendix A for the parameterization). The job finding rate is given by:

$$\theta_e(t|x, \tau_0, s, v_e) = \phi_e(t)\psi_e(\tau_0 + t) \exp\Big[x\beta_e + \delta_{t-s} I(t > s)\Big] v_e \tag{5}$$

Estimation of equation (5) yields an inconsistent estimate of the treatment effects if program participation is (conditional on the observed characteristics) non-random. To account for this, program participation is modeled jointly, also using a mixed

[19] 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). See Lombardi et al. (2018) for a more extensive list of papers that apply the model for other purposes than evaluating active labor market policies.

Table 3: Hazard ratio estimates of the treatment parameters in the 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:						
$\delta_{(1-3\text{ months})}$	0.874	0.014	0.899	0.045	0.834	0.019
$\delta_{(4-6\text{ months})}$	1.036	0.016	0.914	0.048	0.987	0.023
$\delta_{(>6\text{ months})}$	1.247	0.015	1.223	0.047	1.181	0.028
Observations	582,580		112,678		428,160	

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).

proportional hazard rate:

$$\theta_p(s|x,\tau_0,v_p) \quad = \quad \phi_p(s)\psi_p(\tau_0+s)\exp(x\beta_p)v_p \tag{6}$$

with all notation similar to equation (5), 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 piece-wise constant function. Estimation of the parameters is performed by maximizing the log-likelihood, in which right-censoring is straightforwardly taken into account.

Again, we first estimate the model using the full sample, and then turn to using the smaller discontinuity sample and the pre-discontinuity sample.[20] The first column of Table 3 shows that program participation is estimated to have a statistically significant negative effect on the job finding rate in the first three months

[20]Full estimation results, including all estimated coefficients, are presented in the appendix in Table C1.

($\delta_{(1-3\text{ months})}$), with the hazard ratio equal to 0.874. In the next three months the effect is very close to zero (hazard ratio is 1.036). After six months ($\delta_{(\geq 6\text{ months})}$) program participation has a significantly positive effect on the probability of finding a job (1.247). 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 couple of months, and a positive effect afterwards (see column (2) in Table 3). Standard errors are larger due to the smaller sample size. Finally, column (3) shows again very similar results based on the pre-discontinuity sample. The choice of sample thus only causes marginal differences in the estimated treatment parameters. All estimation results show evidence for correlated unobserved heterogeneity, expressed by two distinct mass points.[21] This provides evidence for the presence of unobserved factors, in addition to the observed variables, affecting both job finding and beginning a program.

The estimates for the parameters δ provide a multiplicative effect on the job finding rates. They 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

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

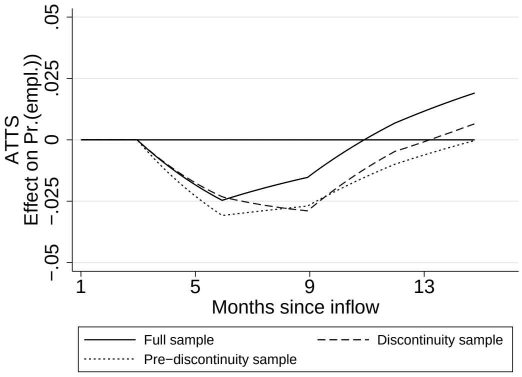
To translate this into the average treatment effect on the treated survivors, we should condition on the rate of receiving treatment after s periods. Therefore, we use the hazard rate model for entering the program, which gives

$$\mathrm{ATTS}(s,t) = \sum_i \frac{\int_v f(s|x_i,v_p)\mathrm{E}[Y^*_{1,t}(s) - Y^*_{0,t}|T > s; x_i, v_e]dG(v_e,v_p)}{\sum_i \int_v f(s|x_i,v_p)dG(v_e,v_p)} \qquad (7)$$

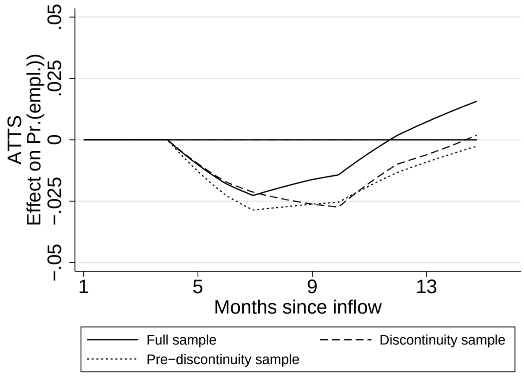
where $f(s|x_i,v_e,v_p) = \theta_p(s|x_i,v_p)\exp(-\int_0^s \theta_e(z|x_i,v_e,s) + \theta_p(z|x_i,v_p)dz)$ is the rate

[21] Note that a simulation study by Lombardi et al. (2018) finds that two mass-points are sufficient to eliminate most selection bias in a timing-of-events model.

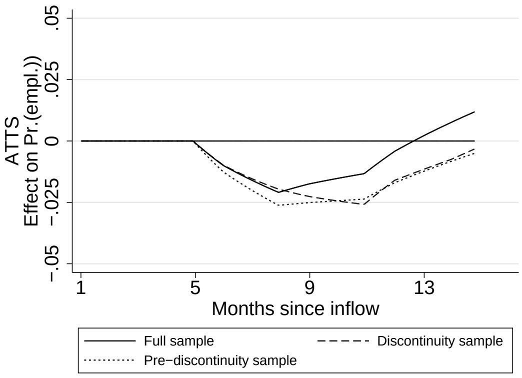
Figure 8: Dynamic matching ATTS estimates based on timing-of-events model with different samples



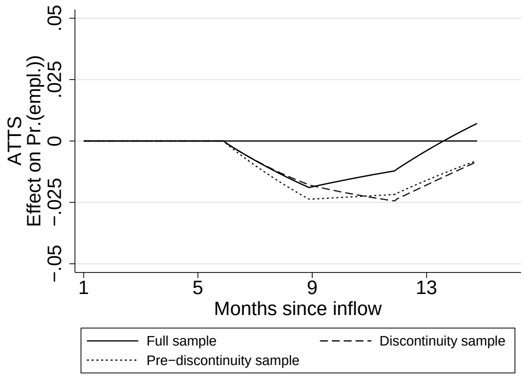
(a) s = 3



(b) s = 4



(c) s = 5



(d) s = 6

at which individual i enters the job search assistance program after s periods. Since we have estimated all right-hand-side parameters in (7), we can obtain an estimate for the ATTS(s, t).

We present estimates for the ATTS for $s = 3, 4, 5, 6$ in Figure 8.[22] Note that these are the same values of s as shown for the dynamic matching estimator in Figure 7. The panels show that the probability of being employed decreases after starting a program, irrespective of the value of s. The magnitude of the negative impact is only slightly smaller though for later treatment ($s = 6$) compared to earlier treatment ($s = 3$). The different samples lead again to similar conclusions. The full sample shows somewhat more optimistic effects but the difference with the other

[22]For computational reasons we compute the ATTS using a random subsample of 1,000 individuals.

samples is small and does not exceed the confidence intervals (see Figures C8, C9 and C10 in the appendix, confidence intervals computed using the delta-method). Finally, the overall magnitude of the negative impact in each of the four panels is smaller compared to what we found using the dynamic matching estimator.

6 Exploiting the policy discontinuity

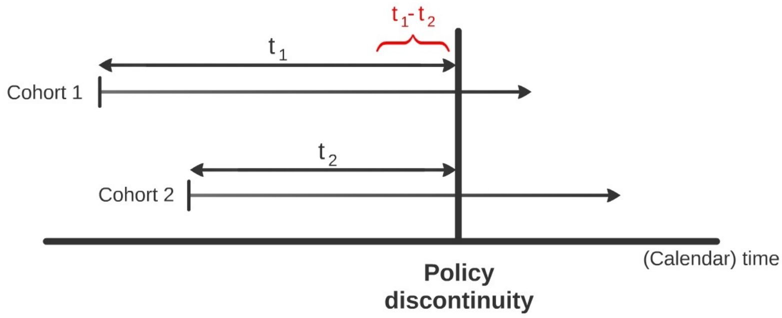
In this section we focus on the policy discontinuity. We first discuss how the exogenous variation due to the policy discontinuity allows to identify the effect of the program on outflow to work. Next we discuss the estimation results.

6.1 Identification

The policy discontinuity is due to the unexpected stop in enrolling new unemployed workers in the external programs. This provides exogenous variation in program participation over inflow cohorts in unemployment. Consider two cohorts of entrants in unemployment. The first enters unemployment at t_1 periods before the policy change. The second cohort enters unemployment later, but still before 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 cohorts to program participation, three conditions should hold. First, the policy discontinuity should be unanticipated by unemployed workers and caseworkers. It should be ruled out that job seekers and caseworkers changed their behavior in anticipation to the policy change

Figure 9: Treatment effect identification



in the period just before March 2010. Our policy change has the advantage that the UI administration only realized late that the budget for these programs had run out 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 unexpectedly rejected.

The second condition is that there should not be compositional differences between cohorts in terms of unobserved characteristics. This condition is equivalent to a conditional independence assumption, but has milder implications. The composition of cohorts is less likely to suffer from selectivity than program assignment. We use weighting to make each cohort equivalent to the March 2010 cohort in composition of observed characteristics. We distinguish 288 groups based on interacting covariates.[23] Define the share of group g in cohort c by $\alpha_{c,g}$. The weight assigned to an observation belonging to group g in cohort c is given by:

$$w_{c,g} = \frac{\alpha_{\text{March2010},g}}{\alpha_{c,g}}$$

We define the survivor functions that will be estimated in the analysis, as the

[23] 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, three education categories and an indicator for being part-time unemployed (less than 34 hours per week.

weighted average of the survivor functions of each cohort-group:

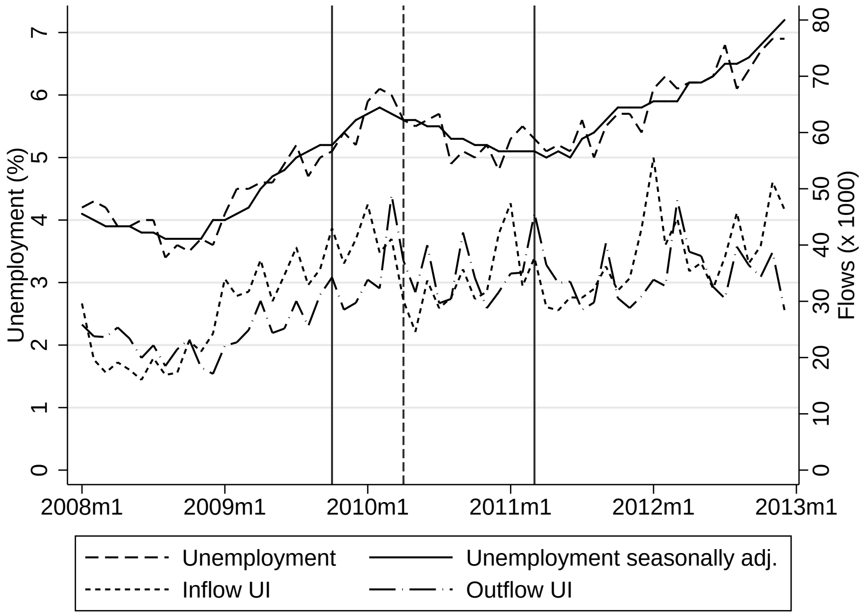
$$S_c(t) = \sum_g w_{c,g} S_{c,g}(t) \tag{8}$$

The third condition is that there should be no differences between the two cohorts in factors that affect job finding, other than the difference in program assignment. However, job finding probabilities change over the business cycle and may fluctuate due to seasonality. To reduce the impact of the business cycle and seasonality, cohorts that are close in time should be compared. But evaluating the program requires cohorts that are substantially spaced apart so that there is sufficient difference is program participation between the cohorts.

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. Changing labor market conditions may affect outcomes in two ways. First, they affect the composition of the inflow into unemployment and this affects aggregate outflow probabilities (which we address by reweighting). Second, labor market conditions affect outflow probabilities directly, as it is more difficult to find employment when unemployment is high (e.g. Van den Berg and Van der Klaauw (2001)).

To deal with business cycle and seasonal effects, we consider a model with an additive structure. Let cohort τ be defined by the month of inflow into unemployment. The survivor function in unemployment (S) has some baseline shape $\lambda(t)$ and further depends on the effect of the business cycle ($b_\tau(t)$), the effect of seasonalities ($l_\tau(t)$) and the effect of entering the program after s periods on being employed

Figure 10: Labor market indicators



Source: Statistics Netherlands, Statline.

after t periods (which is the ATTS(t, s)). Furthermore, $f(s)$ describes the density function of program enrollment after s periods. Our additive specification of the survivor function is given by

$$S_\tau(t) = \lambda(t) + b_\tau(t) + l_\tau(t) + \int_{s=0}^{t} \text{ATTS}(s,t) f(s) \tag{9}$$

The policy discontinuity caused a stop on program entry at moment $\bar{\tau}$, without affecting $f(s)$ prior to this moment. Comparing the survivor function of two cohorts that entered unemployment before the policy discontinuity gives

$$\begin{aligned} S_\tau(t) - S_{\tau'}(t) &= b_\tau(t) - b_{\tau'}(t) + l_\tau(t) - l_{\tau'}(t) + \\ & \int_{s=0}^{\bar{\tau}-\tau} \text{ATTS}(s,t) f(s) - \int_{s=0}^{\bar{\tau}-\tau'} \text{ATTS}(t,s) f(s) \end{aligned} \tag{10}$$

Assuming that τ' describes the more recent inflow cohort ($\tau' > \tau$), then

$$\int_{s=0}^{\bar{\tau}-\tau} \text{ATTS}(s,t) f(s) - \int_{s=0}^{\bar{\tau}-\tau'} \text{ATTS}(s,t) f(s) = \int_{s=\bar{\tau}-\tau'}^{\bar{\tau}-\tau} \text{ATTS}(s,t) f(s)$$

Thus, this simple cohort difference identifies the impact of program participation plus potential business cycle and seasonality differences. To eliminate the effect of seasonality we consider the inflow cohorts 12 months earlier, so $\tau - 12$ and $\tau' - 12$. We take t such that both these inflow cohorts are up to t unaffected by the policy discontinuity, then

$$\Big[S_\tau(t) - S_{\tau'}(t)\Big] - \Big[S_{\tau-12}(t) - S_{\tau'-12}(t)\Big] = b_\tau(t) - b_{\tau'}(t) - b_{\tau-12}(t) + b_{\tau'-12}(t) + \int_{s=\bar{\tau}-\tau'}^{\bar{\tau}-\tau} \gamma(s,t) f(s) \quad (11)$$

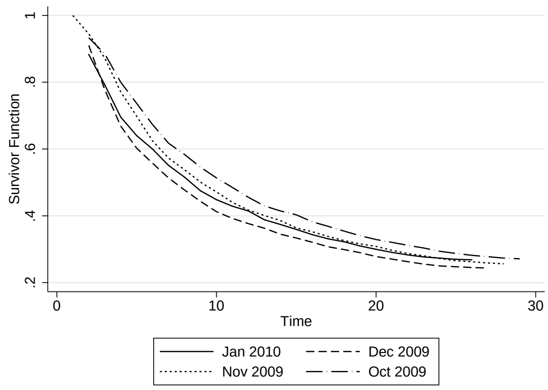
This double difference identifies the impact of program participation if we make the identifying assumption $b_\tau(t) - b_{\tau-12}(t) = b_{\tau'}(t) - b_{\tau'-12}(t)$. This assumption imposes that during the observation period the business cycle effects change only very smoothly. Recall that Figure 10 suggests that, if the observation period is sufficiently small, seasonal effects are much larger than business cycle effects. Our double-difference 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. In Appendix B we discuss in more detail the assumption that business cycle effects are small (which is similar to the standard common trend assumption in difference-in-differences).

In our empirical approach we estimate the intention-to-treat effects, specified as $\int_{s=\bar{\tau}-\tau'}^{\bar{\tau}-\tau} \mathrm{ATTS}(s,t) f(s)$, but using the empirical program participation rate $f(s)$ it is straightforward to obtain $\mathrm{ATTS}(s,t)$, which describes the average treatment effect on the treated survivors.

6.2 Estimation results

We start by defining which cohorts to compare. The time interval between cohorts should be small to minimize business cycle effects, but the trade-off is that more time between cohorts increases the difference in exposure to program participation.

Figure 11: Survivor functions by month of inflow



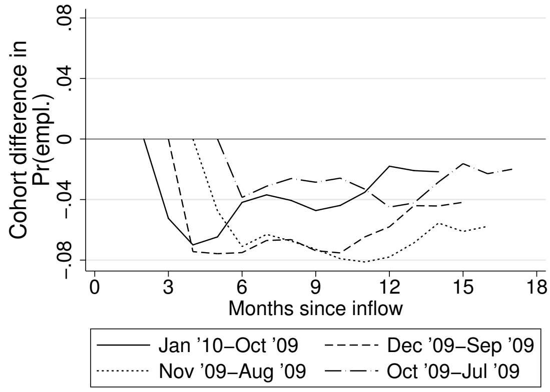
We use cohorts three months apart. To exploit the policy discontinuity, the cohorts should not enter unemployment too long before March 2010. Therefore, we use the cohorts of October 2009 until January 2010, facing between five and two months of potential program participation, respectively. Each cohort is compared to the cohort entering unemployment three months earlier. The survivor function of each cohort is presented in Figure 11, showing that around 50% of the UI recipients find work within 12 months.

We first take the difference between the survivor function and the survivor function of the cohort entering unemployment three months earlier.[24] This compares the outflow in a cohort without enrollment in the program to the outflow in a cohort with regular enrollment in the program during the period before the policy discontinuity. We condition on survival and no-treatment up to the duration at which the later cohort reaches the policy discontinuity. The three-months differences are presented in panel (a) of Figure 12. We find very similar patterns across the different comparisons: a negative effect on job finding ranging from 3%-points to almost 8%-points which decreases in magnitude over time. The negative effects persist up to at least 13 to 17 months of unemployment.

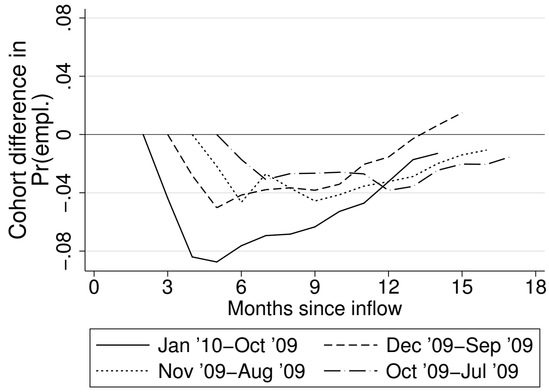
These estimates are based on simple differences between cohorts (equation (10)). By subtracting the same differences from a year earlier, we correct for seasonalities

[24]All estimates presented in this section are estimated using weights as discussed in Subsection 6.1.

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



(a) Simple difference



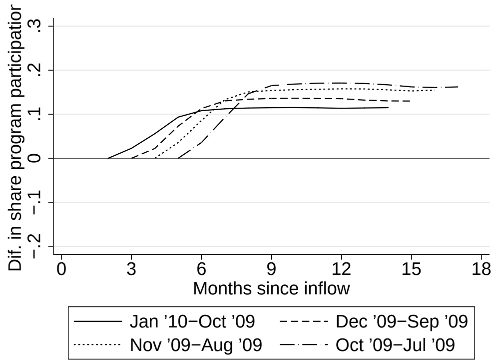
(b) Double difference

(equation (11)). Estimates from such a "difference-in-differences" approach are presented in panel (b) of Figure 12.[25] We find a very similar negative effect on job finding in the first months, although the reductions at longer durations are greater in magnitude than in the single difference estimates. At longer durations the estimates converge to a zero effect (or even a slightly positive impact). The effects are statistically significant for about the first 15 months (see Figures C11 and C12 in the appendix for 95% confidence intervals). Each comparison measures the effect of 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.

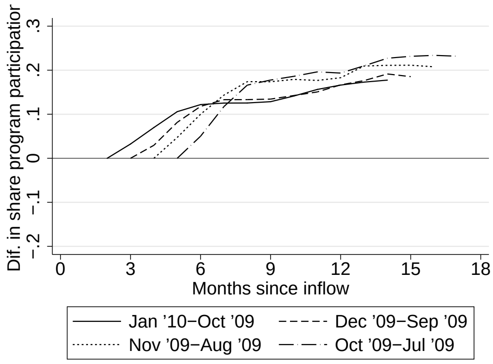
Thus, we find a negative impact of program participation on job finding that is consistent across different cohort comparisons and across the two estimators. This finding is in line with the lock-in effect that we have discussed earlier. These estimates measure intention-to-treat effects, and should be divided by the cohort differences in treatment assignment to obtain average treatment effects.

[25]When estimating the effects 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

We estimate the difference in program participation by

$$\int_{s=0}^{\bar{\tau}-\tau'} \hat{f}_{\tau'}(s)ds - \int_{s=0}^{\bar{\tau}-\tau} \hat{f}_{\tau}(s)ds \tag{12}$$

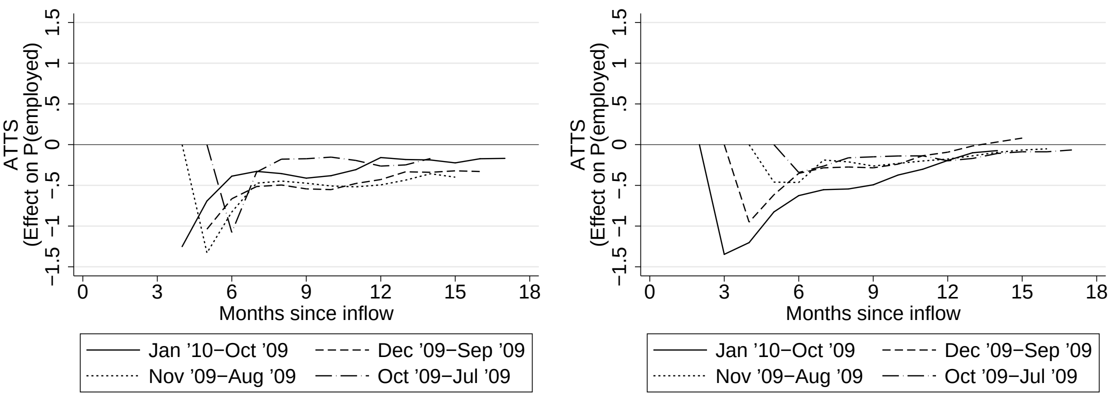
where $\bar{\tau}$ is the moment of the policy discontinuity (March 2010) and τ' and τ are two inflow cohorts, e.g. January 2010 and October 2009. The functions $\hat{f}$ describe the empirical program participation rates in these inflow cohorts. Panel (a) of Figure 13 shows estimated differences in program participation rates. There is a clear increase at the moment the first cohort reaches March 2010. The difference increases for approximately three months to about 10% to 20%-points. After that the comparison cohort reaches March 2010 and both cohorts no longer enroll in the program. The difference in program participation can also be computed using the same "difference-in-differences" approach, which are presented in panel (b) of Figure 13. The differences are less smooth, but exhibit the same pattern.[26]

The estimated ATET(s,t) are presented in Figure 14. Since differences in treatment participation between cohorts accrue over a three-months period, the initially small differences inflate the intention-to-treat effects substantially. To facilitate visual presentation, we remove some of the extreme estimates in the early months.[27]

[26] All differences are highly statistically significant (see Figures C13 and C14 in the appendix for confidence intervals).

[27] Complete graphs are provided in the Appendix (Figure C15), as well as confidence intervals for each separate set of estimates (Figures C16 and C17).

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



(a) Simple difference

(b) Double difference

We find that the impact on the probability on being employed is around 20 to 40%-points, both when using simple differences (panel (a)) and double differencing (panel (b)). The effects decline over time. While a small negative impact remains in the simple-difference estimates, the impact seems to disappear in the double-differencing estimates after about 15 months.

Note that from an identification perspective it is a major advantage that participation was reduced to zero (rather than to some positive value), because that implies that our estimates can be interpreted as average treatment effects on the treated, rather than local average treatment effects.[28]

7 Discussion

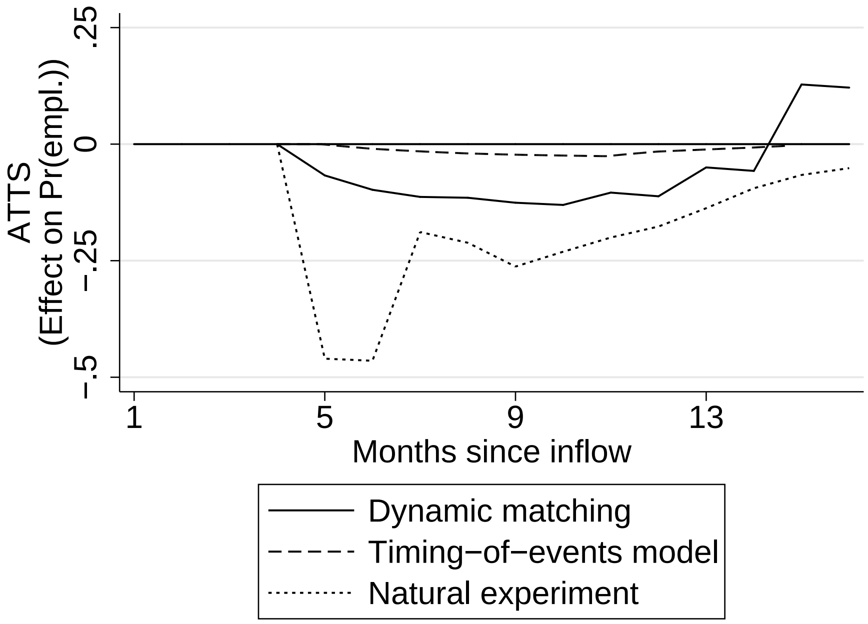
We have estimated the impact of the (private-sector provided) activation programs on finding employment using three approaches, (*i*) a dynamic matching estimator, (*ii*) the timing-of-events model, and (*iii*) exploiting the policy discontinuity as a natural experiment. In this section we compare the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

Recall from the previous sections that the estimated impact of the programs is

[28]In the terminology of instrumental variables, there is full compliance with the instrument (the discontinuity) and there are no "always-takers".

almost unaffected by the starting duration. For ease of the discussion, we therefore consider for each method the estimated impact of program participation starting at $s = 5$, which is the unemployment duration at which actual enrollment in the program is highest (see Figure C2 in the appendix).[29] The three methods estimate the same treatment effect on the treated survivors.

Figure 15: Comparing different methods (for ATTS)



Note: The dynamic matching estimate corresponds to the discontinuity sample line in panel (c) of Figure 7. The timing-of-events estimate corresponds to the discontinuity sample line in panel (c) of Figure 8. The quasi-experimental estimate corresponds to the November 2009 - August 2009 comparison in panel (b) of Figure 14.

Figure 15 shows that all three methods have the same trend in the estimated treatment effects but the magnitudes of the estimated treatment effects differ. Immediately after enrolling in the program, the job finding rate declines, and after some months the negative effect becomes smaller. All three methods yield the same conclusion that participation in the program postpones job finding, but in the long-run does not affect the probability of having work.[30] Vikström (2017) finds a similar

[29]We take for the dynamic matching method and the timing-of-events method the estimated effects based on the (smaller) discontinuity sample, to make them most comparable to the quasi-experimental estimates. Finally, for the quasi-experimental estimates we use the double-difference estimates.

[30]We cannot predict what happens in the very long-run beyond our observation period, but job

pattern in his dynamic evaluation of a Swedish work practice program. The policy conclusion from each of the methods is that the programs are not cost-effective.

In their meta-analysis, Card et al. (2011) classify evaluation study results as a positive, negative or zero impact. However, if one goes beyond the sign of the impact, we find substantial differences in the magnitudes of the estimated treatment effects. The quasi-experimental estimate is largest. This approach estimates the average treatment effect by dividing the intention-to-treat effect by a small fraction of treated individuals. This causes that the treatment effect is estimated relatively imprecise and the estimates based on the timing-of-events model and the matching approach fall at most durations in the 95% confidence interval of the quasi-experimental approach (see panel (c) of Figure C17, in the Appendix).[31]

The difference between the magnitudes of the estimates may suggest that not all identifying assumptions hold. First, the timing-of-events study finds some relevant unobserved heterogeneity. This may either imply that the conditional independence assumption required for the matching estimator might be violated or that the mixed proportional hazard specification in the timing-of-events model is too restrictive. Second, the common-trend assumption in the quasi-experimental approach could be violated, leading to a (small) downward bias. Finally, the no-anticipation assumption is required to define treatment effects. Due to the unexpected policy discontinuity this assumption is less likely to be violated in the quasi-experimental evaluation.

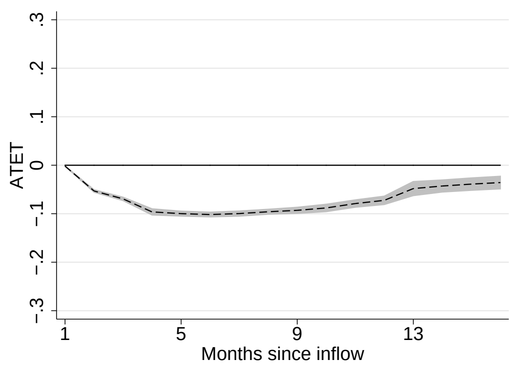
Our empirical analysis provides estimates for the ATTS(t, s), which is the impact on individual employment probabilities for specific values of s and t. A policy maker might be interested in the average effect for all participants. We have defined the ATET(t) in equation (2), which is essentially an average impact of program

search assistance programs typically do not have very strong long-run effects (e.g. Card et al. (2010)).

[31] We acknowledge that since the estimates are based on different models and are not independent, the confidence intervals may not provide a fully precise approach to determine whether the differences between methods are significant.

participation weighted by the inflow in the program at duration s. Since all three identification strategies suggest a negative impact of program participation for any s, the corresponding ATET(t) will also be negative for all values of t. As illustration Figure 16 shows the estimated ATET(t) for the dynamic matching estimates. For any potential unemployment duration t participation in the program reduces the probability to have found work before.

Figure 16: ATET(t), based on matching estimates



8 Conclusion

Several methods are available when evaluating activation programs for unemployed job seekers. In this paper we compare estimates from three different methods. First, we apply a dynamic matching estimator, second we estimate the timing-of-events model, and third we exploit exogenous variation in program participation, caused by budgetary problems of the UI administration.

All three methods suggest a significantly negative effect of program participation on employment. The quasi-experimental estimates suggest reductions in the probability of being employed of up to 50%-points early after inflow in the program, the matching and timing-of-events estimates are smaller in magnitude (2.5 - 15%-points). In the longer run, all three methods suggest an (imprecise) zero effect on employment. The robust conclusion drawn from each approach is that the programs are not effective in increasing outflow, and even reduce job finding rates in the short and medium long run. We have focused on employment as the key outcome variable. We cannot assess whether the program affects job quality (such as the salary or job stability), and thus we cannot exclude that perhaps it has a positive impact along these dimensions.

We find some relevant unobserved heterogeneity in the timing-of-events model which might explain the difference in magnitude compared to the dynamic matching estimates. However, it does not to change the pattern of treatment effects between these methods. Our results concur with the meta-analysis performed by Kluve (2010), who does not find a relation between the methodology and the likelihood of estimating positive or negative effects. Our results confirm that this also holds when evaluating the same program using the same data, rather than when comparing across studies.

These results are based on a large sample that covers all individuals that became unemployed in the Netherlands during a two year period. Thus, results are not specific to particular subgroups but apply to the entire population. However, the results are specific to the program assignment rule. Targeting the program using more selective criteria can change the effectiveness of the program. Our results also concur with other studies showing the lack of effectiveness of job search assistance programs offered by commercial providers (e.g. Behaghel et al. (2014) and Cottier et al. (2015)).

References

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., 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.

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.

Cottier, L., Kempeneers, P., Flückiger, Y., and Lalive, R. (2015). Does outsourcing job search assistance help job seekers find and keep jobs? IZA Discussion Paper.

De Groot, N. and Van der Klaauw, B. (2019). The effects of reducing the entitlement period to unemployment insurance benefits. *Labour Economics*, Forthcoming.

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. 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.

Lombardi, S., Van den Berg, G., and Vikström, J. (2018). Empirical monte carlo evidence on estimation of timing-of-events models. Technical report, mimeo.

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.

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. and Van der Klaauw, B. (2001). Combining micro and macro unemployment duration data. *Journal of Econometrics*, 102(2):271–309.

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. (2017). Dynamic treatment assignment and evaluation of active labor market policies. *Labour Economics*, 49(C):42–54.

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:

$$\theta_e(t|x, \tau_0, s, v_e) = \phi_e(t)\psi_e(\tau_0 + t) \exp\Big[x\beta_e + \delta_{t-s} I(t > s)\Big] v_e \tag{13}$$

and the entry rate into the program by:

$$\theta_p(s|x, \tau_0, v_p) \quad = \quad \phi_p(s)\psi_p(\tau_0 + s) \exp(x\beta_p) v_p \tag{14}$$

We specify the duration dependence patterns $\phi_e(t)$ and $\phi_p(s)$ as piecewise constant, so $\phi_j(t) = \exp\left(\sum_{m=1}^{M} \pi_{jm} I_m(t)\right)$ 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

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

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.

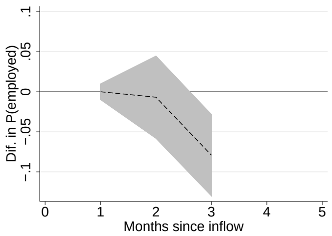
B Common trend assumption

Exploiting the policy discontinuity requires a common trend assumption. Below we discuss the justification of this assumption.

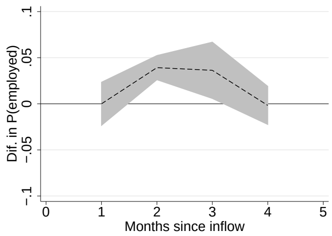
The assumption that the remaining terms in equations (10) and (11) are negligible has some similarities with the common trend assumption in a difference-in-differences estimator. The simple difference estimator require that in the absence of the policy discontinuity the employment rate of the January 2010 cohort would have been the same as the employment rate of the October 2009 cohort (and similar for the other comparisons). Similarly, the double difference estimator requires that in the absence of the policy discontinuity, the *difference* in employment rate between the January and October cohort would be 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 assumptions, by investigating the survivor functions over the first months of each cohort, so before the later cohort is affected by 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 the survivors differences for $t \leq t_2$, without conditioning on survival up to a certain duration.

Estimates are presented in Figure B1 for single differences (panels (a), (b) and (c)), including 95% confidence intervals computed using bootstrapping. We find that for the single difference estimator, the first comparison (panel (a)) shows a negative and significant difference. On the other hand, the second and third comparisons (panels (b) and (c)) show a significant positive difference that diminishes over time. Although these suggest that the common trends assumption is violated and thus our estimates might be biased, the lack of a consistent pattern across the three comparisons makes us confident that the main conclusion drawn from the estimates

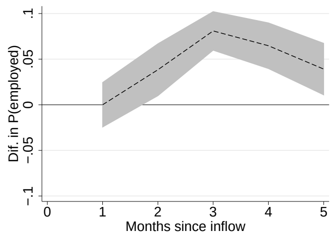
Figure B1: Common trend tests



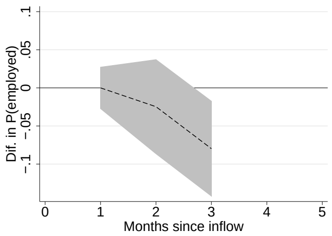
(a) Dec. ’09 - Sep. ’09, simple dif.



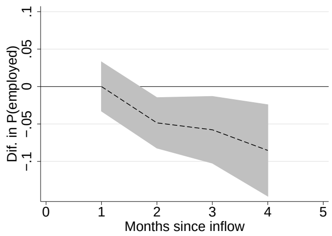
(b) Nov. ’09 - Aug. ’09, simple dif.



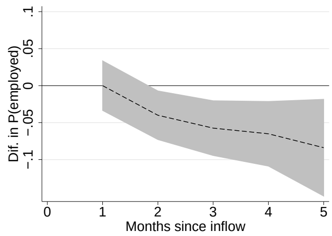
(c) Oct. ’09 - Jul. ’09, simple dif.



(d) Dec. ’09 - Sep. ’09, double dif.



(e) Nov. ’09 - Aug. ’09, double dif.



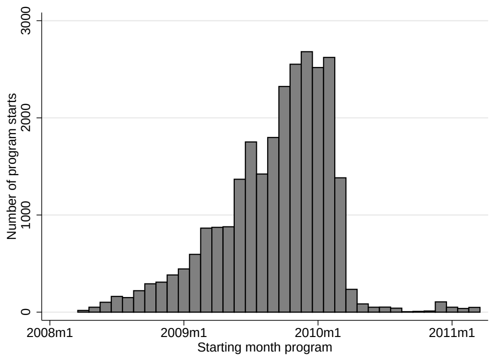
(f) Oct. ’09 - Jul. ’09, double dif.

in Figure 12 still holds.[32] We perform a similar check for the double differences estimator in panels (d), (e) and (f) of Figure B1. Here the patterns are more similar across the comparisons, showing a significant negative difference. As a result, our estimates will be biased downwards somewhat, and thus provide lower bounds of the effect. This might explain that the magnitude of the negative impact seems particularly large in some comparisons. Overall, these results are not too surprising, as our approach can clearly not fully control for all business cycle variation in job finding rates.

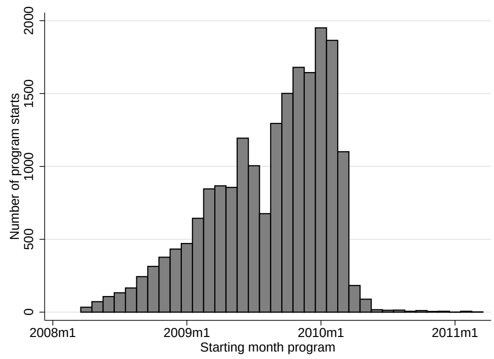
[32]In particular, the three panels suggest that estimates from the December - September comparison are biased downwards, estimates from the November - August comparison are nearly unbiased and estimates from the October - July comparison are biased upwards. Given that all three comparisons yield similar results (see Figure 12), we argue that the bias is unlikely to alter the main conclusions.

C Additional empirical material

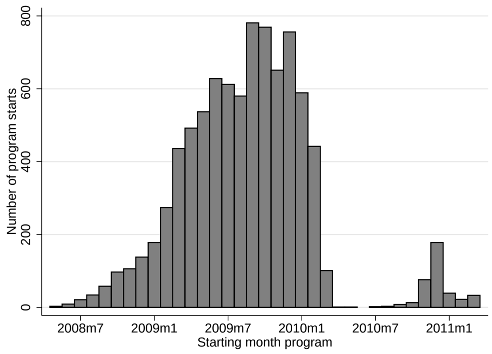
Figure C1: Starting dates of externally provided programs



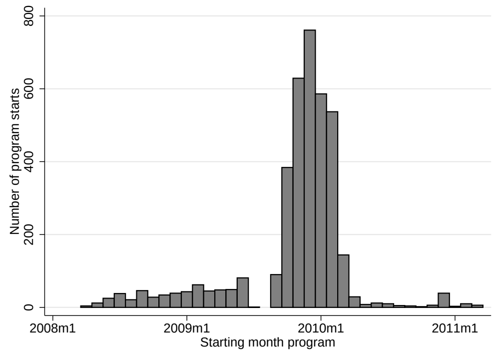
(a) IRO (Individual reintegration agreement)



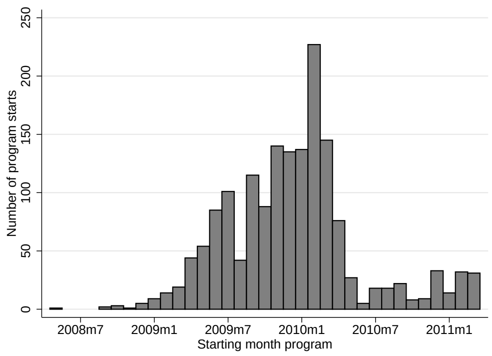
(b) Short training



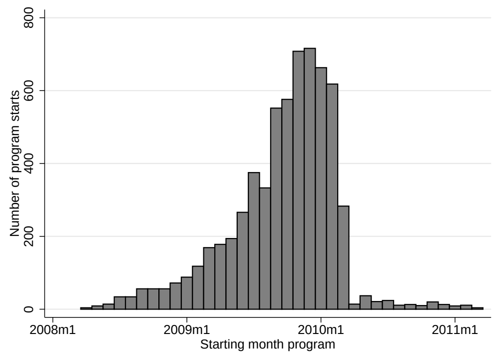
(c) Jobhunting



(d) Standard programs

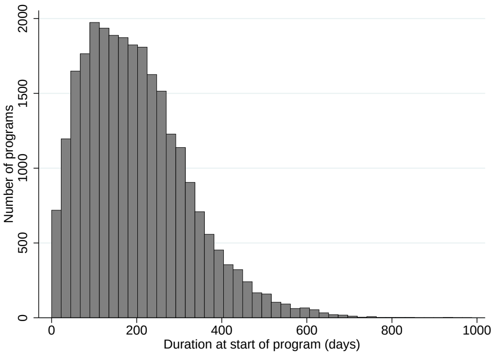


(e) Learn-work jobs

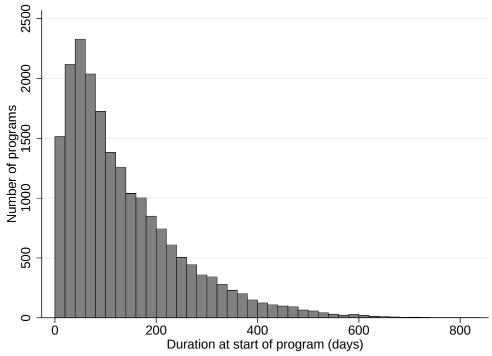


(f) Schooling

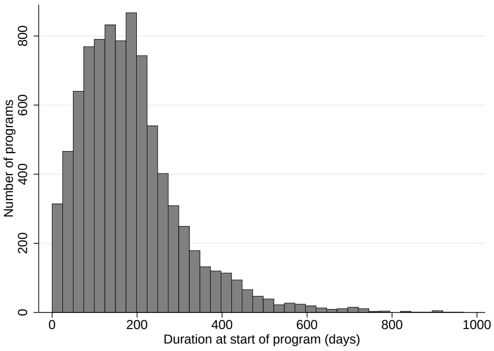
Figure C2: Timing of externally provided programs



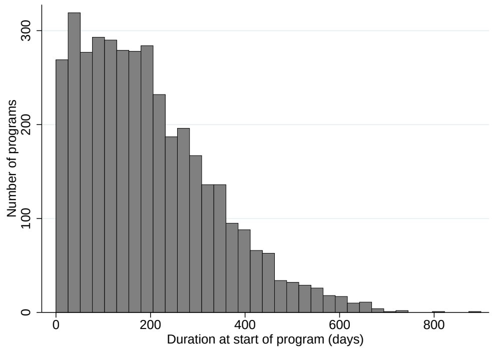
(a) IRO (Individual Reintegration Agreement)



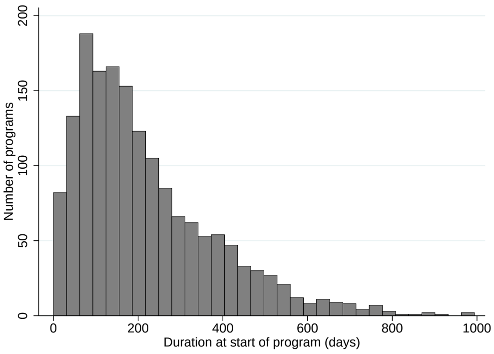
(b) Short training



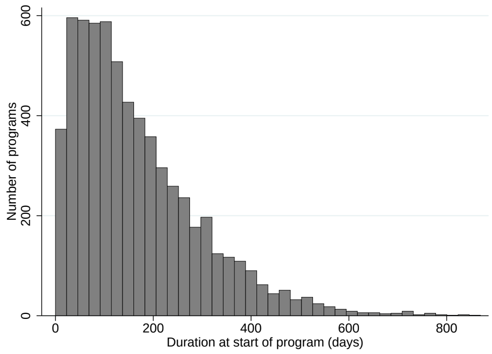
(c) Jobhunting



(d) Regular programs

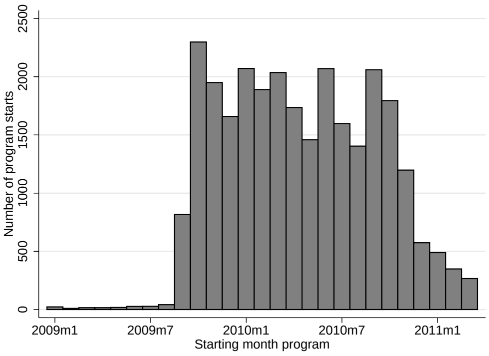


(e) Learn-work jobs

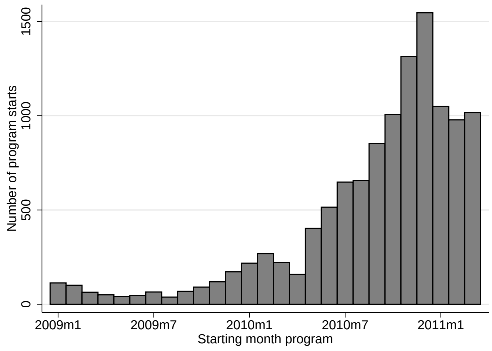


(f) Schooling

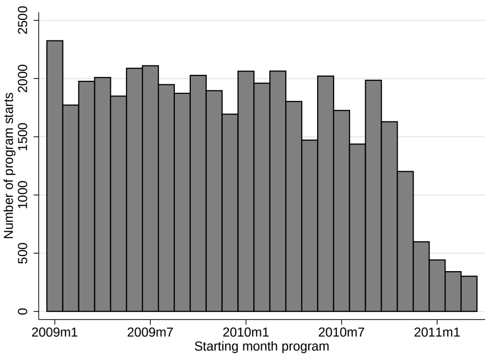
Figure C3: Starting dates of internally provided programs



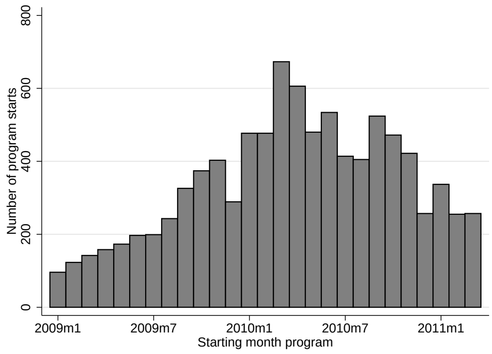
(a) CTC



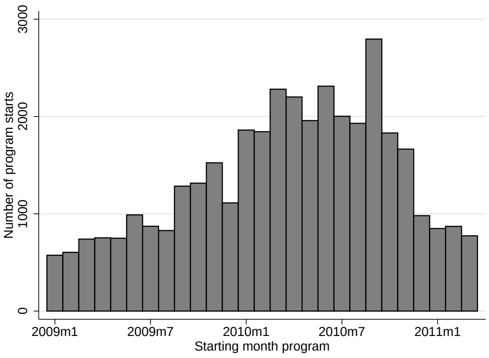
(b) Vacancy referral



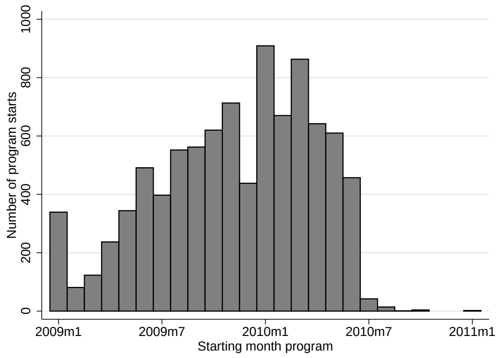
(c) Tests (different types)



(d) Employment on trial basis

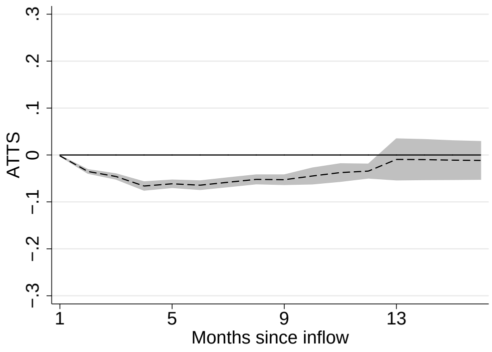


(e) Workshop

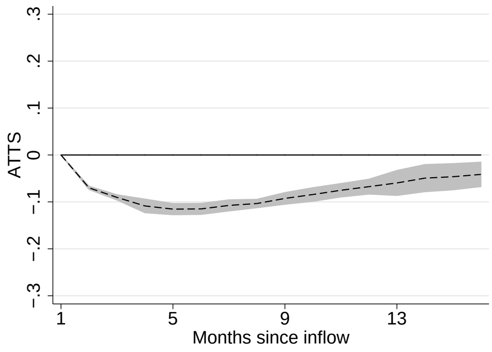


(f) Entrepeneurship support program

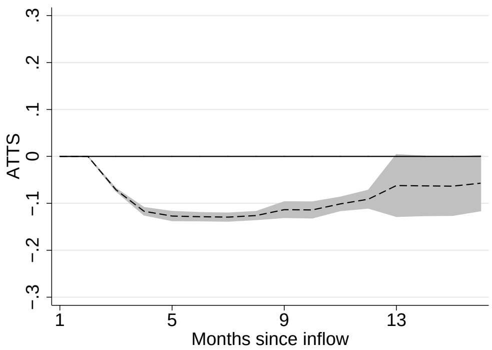
Figure C4: Dynamic matching estimator with 95% confidence intervals (full sample)



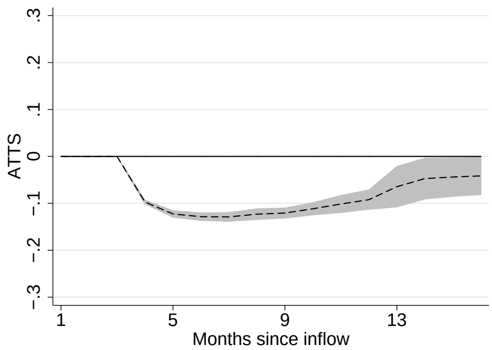
(a) s = 1



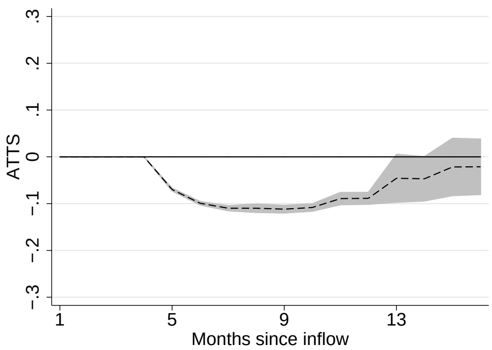
(b) s = 2



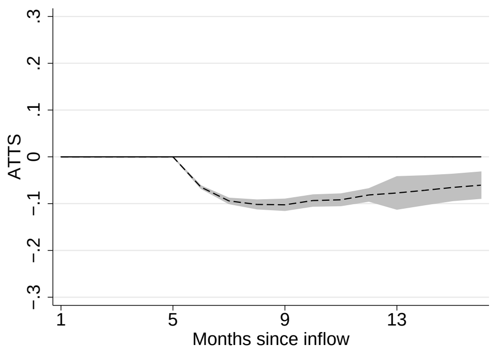
(c) s = 3



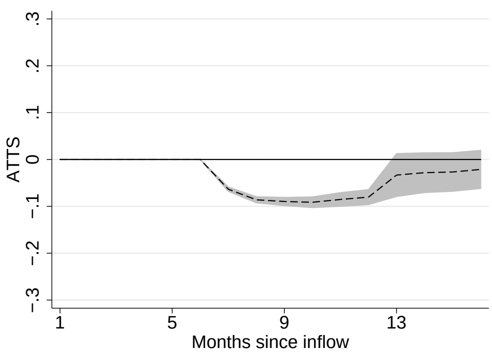
(d) s = 4



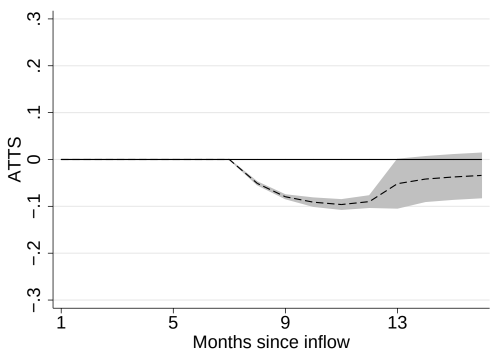
(e) s = 5



(f) s = 6



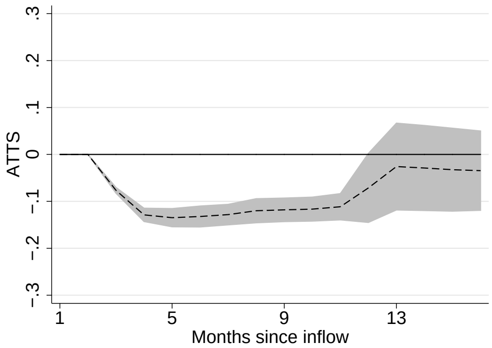
(g) s = 7



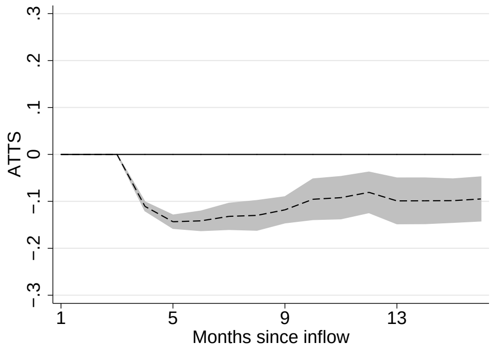
(h) s = 8

Note: The confidence intervals have been constructed using bootstrapping (20 repetitions).

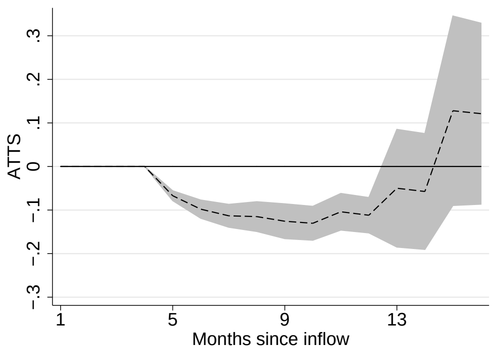
Figure C5: Dynamic matching estimator with 95% confidence intervals (discontinuity sample)



(a) s = 3



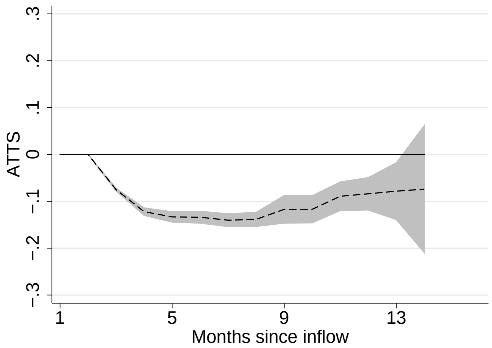
(b) s = 4



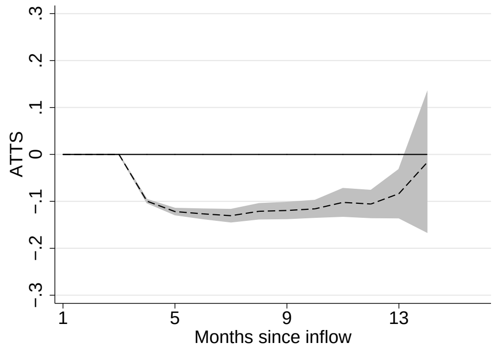
(c) s = 5

Note: The confidence intervals have been constructed using bootstrapping (20 repetitions).

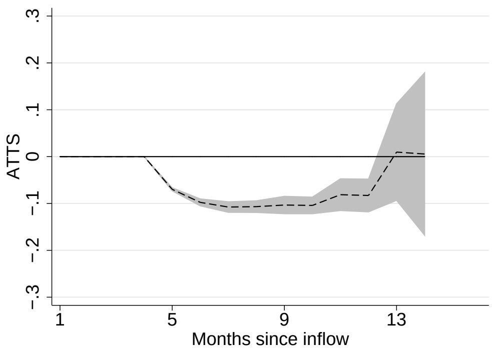
Figure C6: Dynamic matching estimator with 95% confidence intervals (censored sample)



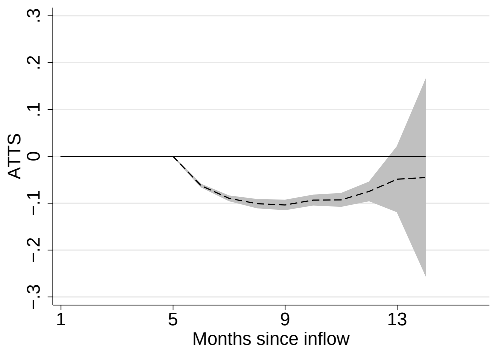
(a) s = 3



(b) s = 4



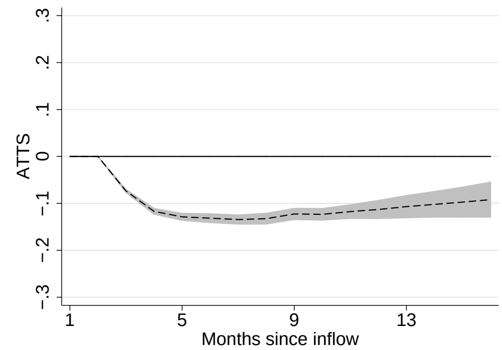
(c) s = 5



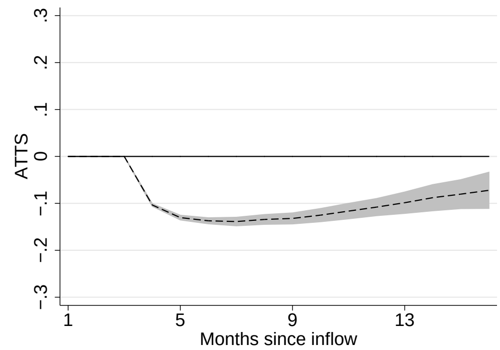
(d) s = 6

Note: The confidence intervals have been constructed using bootstrapping (20 repetitions).

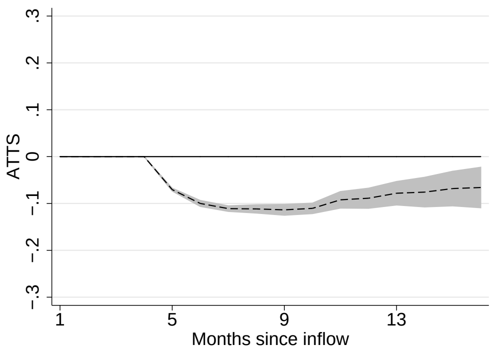
Figure C7: Dynamic matching estimator with 95% confidence intervals (few X-variables)



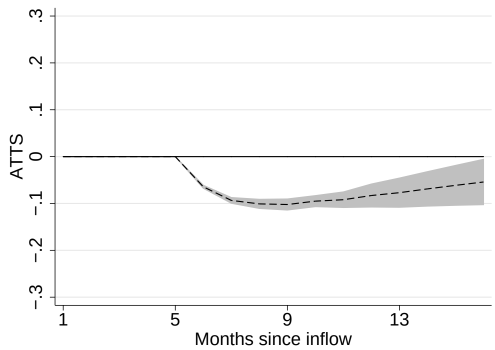
(a) s = 3



(b) s = 4



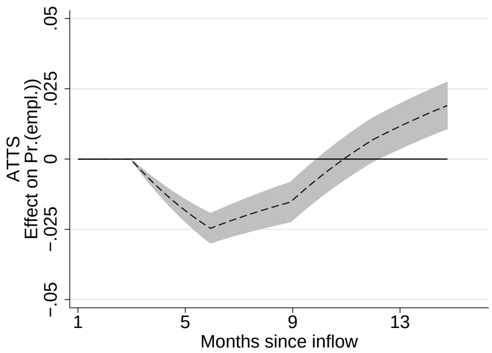
(c) s = 5



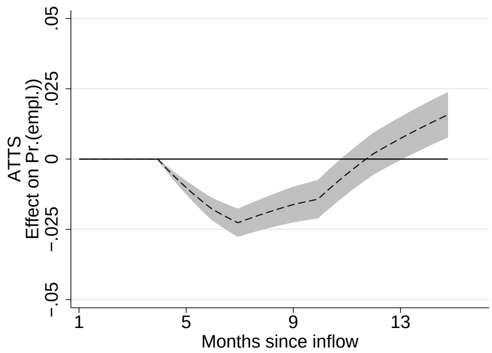
(d) s = 6

Note: The confidence intervals have been constructed using bootstrapping (20 repetitions).

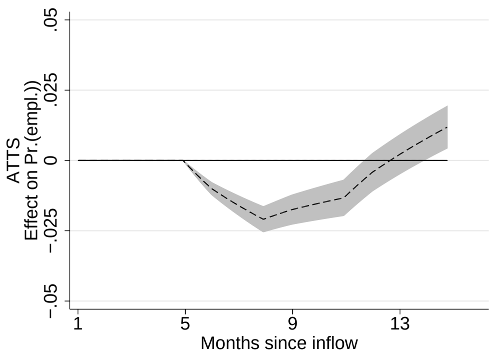
Figure C8: Timing-of-events estimates with 95% confidence intervals (full sample)



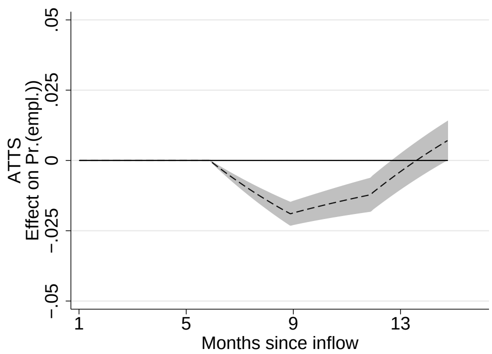
(a) s = 3



(b) s = 4

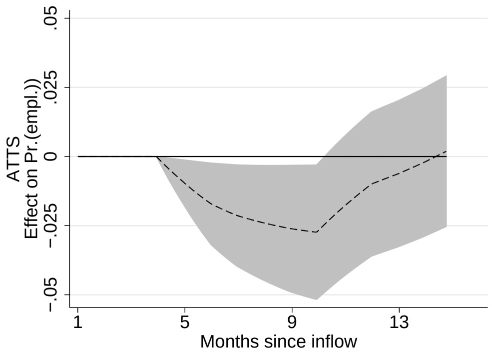


(c) s = 5

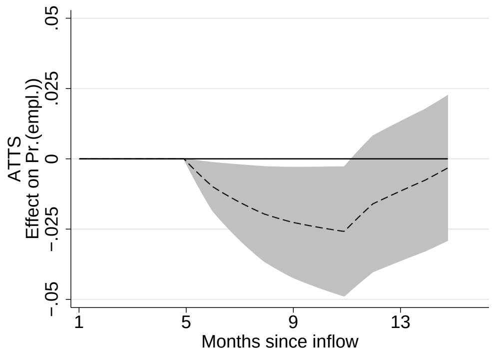


(d) s = 6

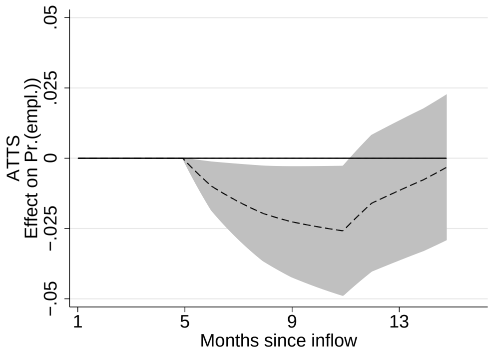
Figure C9: Dynamic matching estimator with 95% confidence intervals (discontinuity sample)



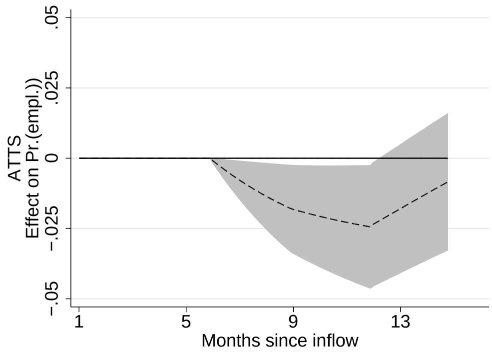
(a) s = 3



(b) s = 4

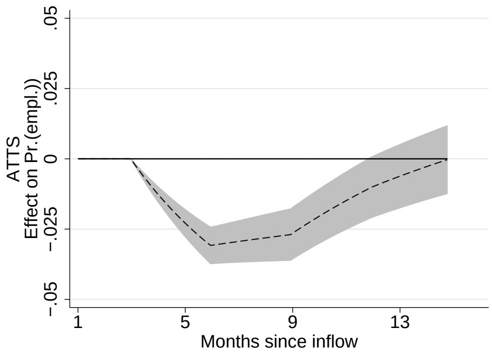


(c) s = 5

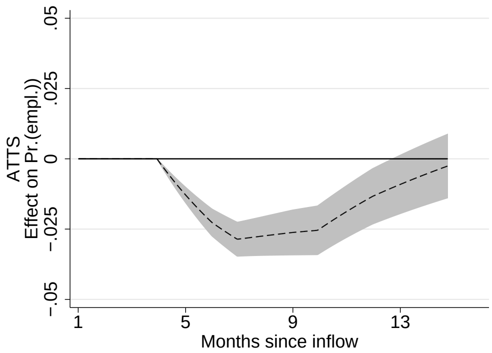


(d) s = 6

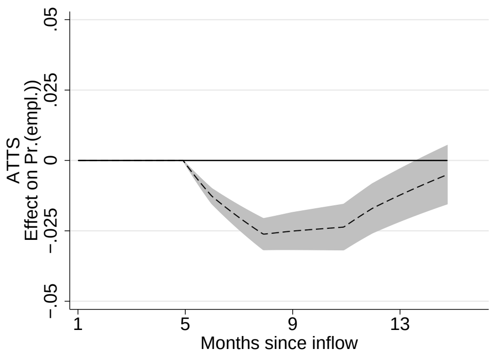
Figure C10: Timing-of-events estimates with 95% confidence intervals (pre-discontinuity sample)



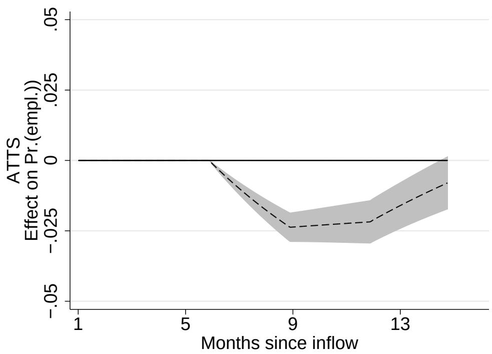
(a) s = 3



(b) s = 4

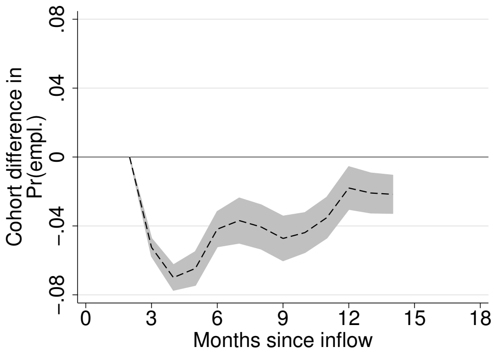


(c) s = 5

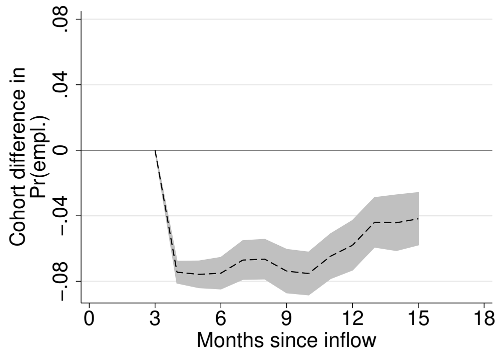


(d) s = 6

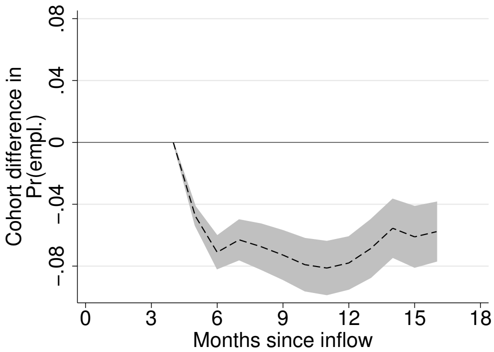
Figure C11: Experimental estimates (single differences) with 95% confidence intervals



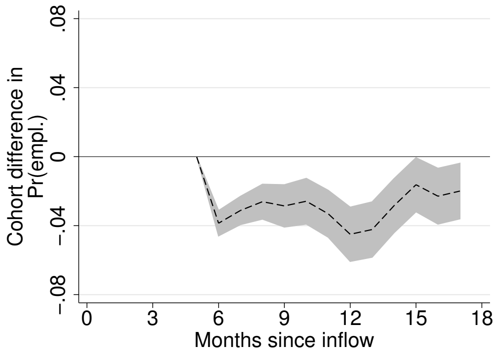
(a) Jan 2010-Oct 2009



(b) Dec 2009-Sep 2009

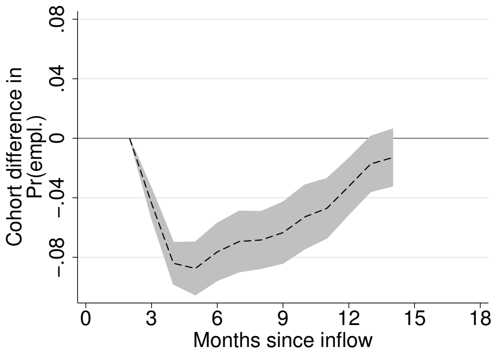


(c) Nov 2009-Aug 2009

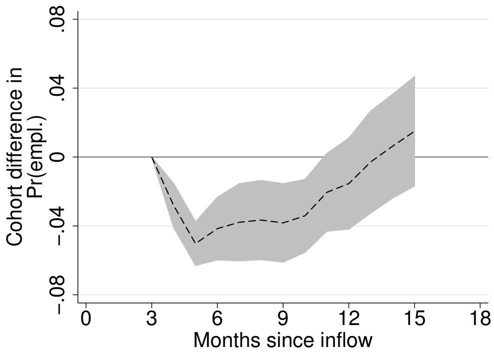


(d) Oct 2009-Jul 2009

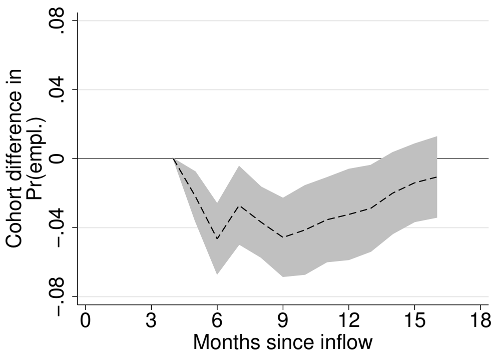
Figure C12: Experimental estimates (double differences) with 95% confidence intervals



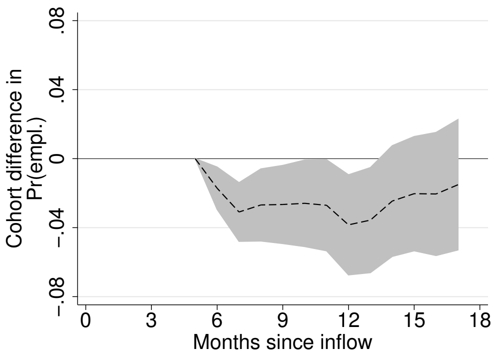
(a) Jan 2010-Oct 2009



(b) Dec 2009-Sep 2009

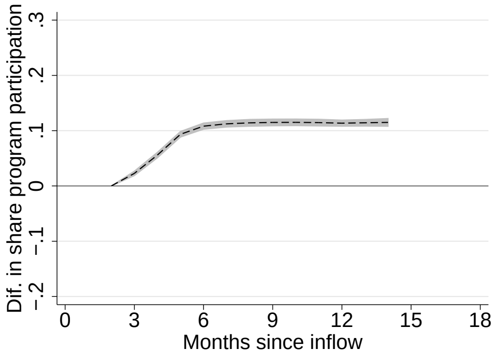


(c) Nov 2009-Aug 2009

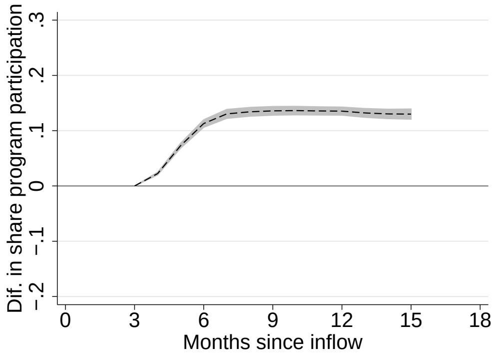


(d) Oct 2009-Jul 2009

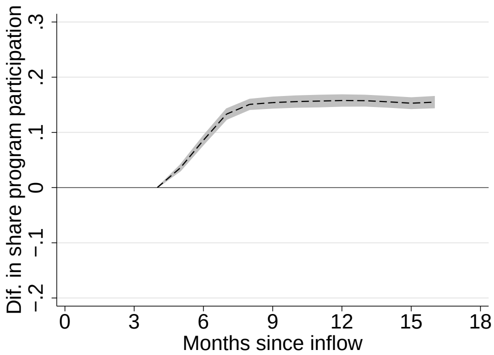
Figure C13: Single difference in treatment share with 95% confidence intervals



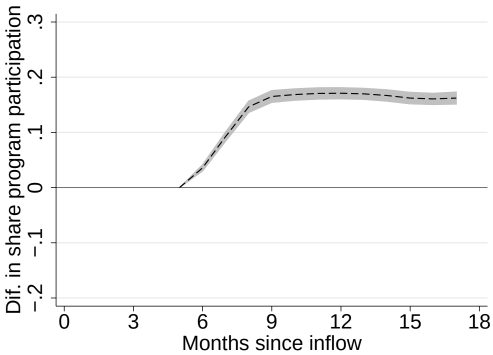
(a) Jan2010-Oct2009



(b) Dec2009-Sep2009

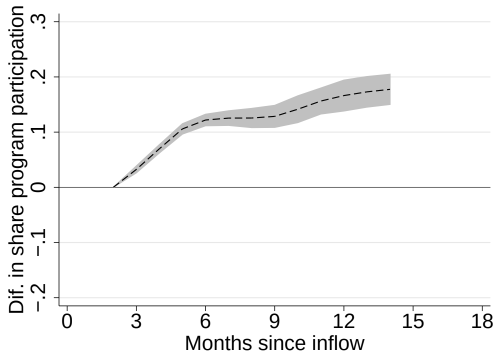


(c) Nov2009-Aug2009

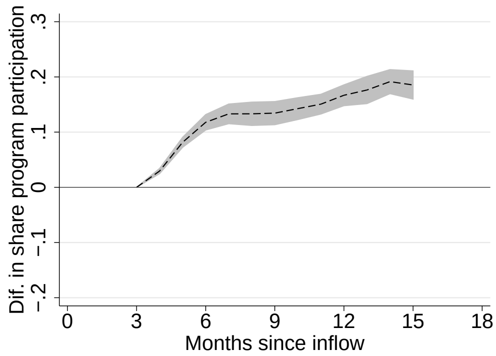


(d) Oct2009-Jul2009

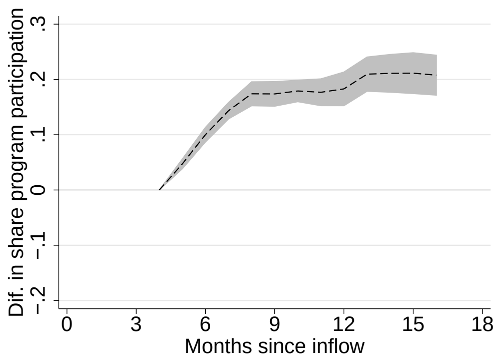
Figure C14: Double difference in treatment share with 95% confidence intervals



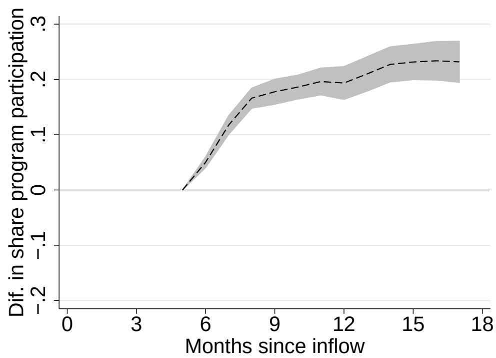
(a) Jan2010-Oct2009



(b) Dec2009-Sep2009



(c) Nov2009-Aug2009



(d) Oct2009-Jul2009

Figure C15: Average treatment effect (including extreme values), single differences

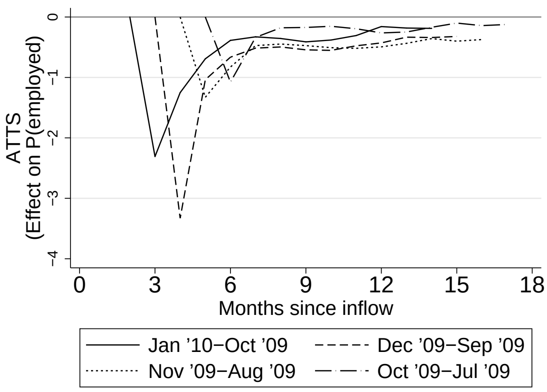
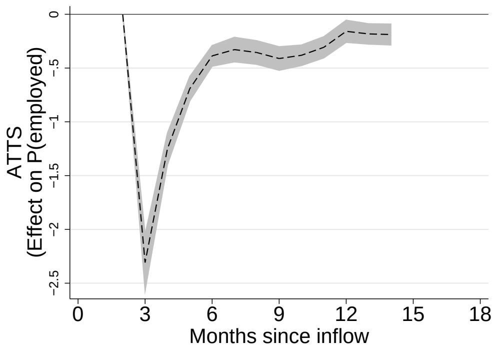
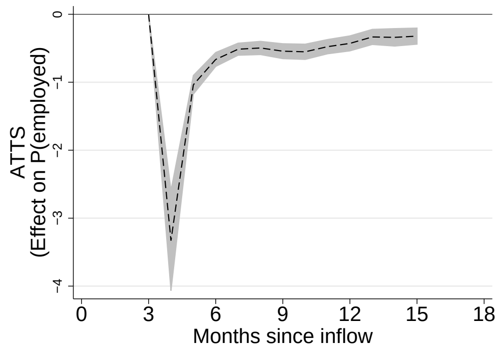


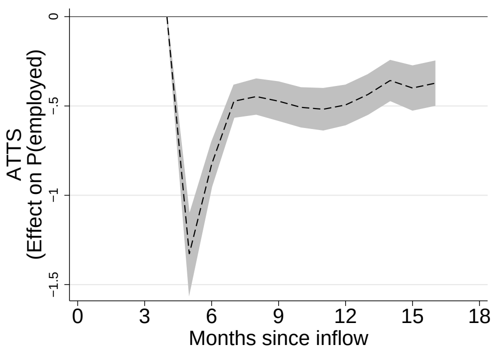
Figure C16: Average treatment effects (single differences) with 95% confidence intervals



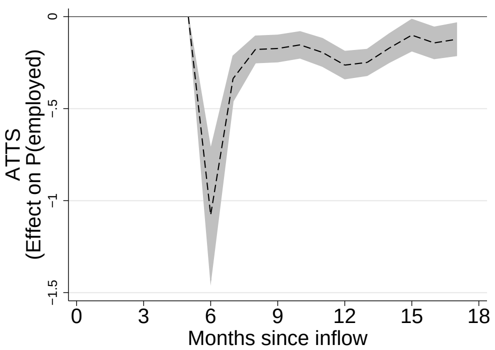
(a) Jan2010-Oct2009



(b) Dec2009-Sep2009

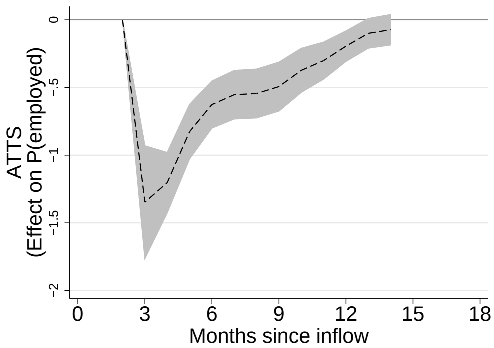


(c) Nov2009-Aug2009

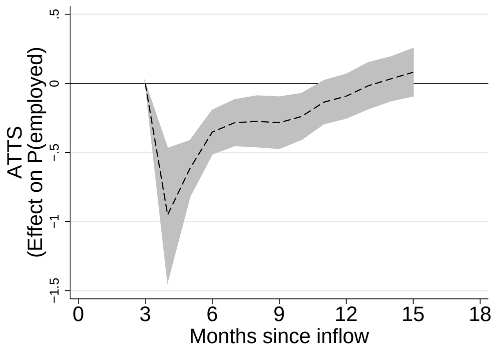


(d) Oct2009-Jul2009

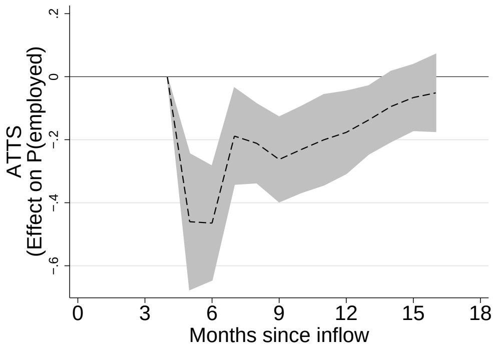
Figure C17: Average treatment effects (double differences) with 95% confidence intervals



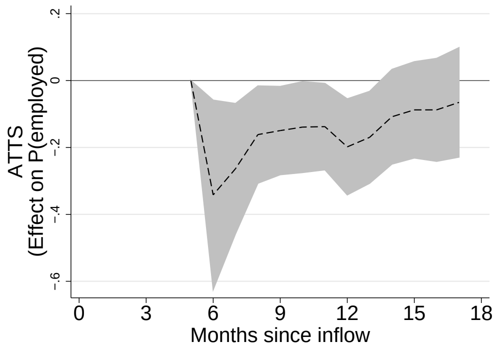
(a) Jan2010-Oct2009



(b) Dec2009-Sep2009



(c) Nov2009-Aug2009



(d) Oct2009-Jul2009

Table C1: 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	0.874	0.014			0.899	0.045			0.834	0.019		
Months 4-6	1.036	0.016			0.914	0.048			0.987	0.023		
Months 7-	1.247	0.015			1.223	0.047			1.181	0.028		
Individual characteristics:												
Female	0.973	0.004	1.045	0.011	0.940	0.010	1.062	0.025	1.049	0.006	1.045	0.011
Age 25-25	0.759	0.007	1.398	0.022	0.773	0.016	1.202	0.042	0.763	0.010	1.437	0.024
Age 35-45	0.645	0.009	1.625	0.024	0.655	0.019	1.329	0.047	0.665	0.012	1.672	0.026
Age 45-55	0.552	0.010	1.837	0.026	0.579	0.023	1.443	0.052	0.607	0.014	1.915	0.028
Age 55-60	0.298	0.012	1.548	0.029	0.330	0.028	1.257	0.060	0.382	0.018	1.613	0.031
Middle educated	1.134	0.005	1.197	0.011	1.103	0.010	1.238	0.024	1.152	0.007	1.193	0.011
High educated	1.230	0.006	1.064	0.014	1.171	0.014	1.069	0.034	1.286	0.009	1.040	0.015
Income (cat. 2)	1.092	0.007	1.173	0.016	1.079	0.015	1.228	0.036	1.087	0.009	1.175	0.016
Income (cat. 3)	1.204	0.007	1.205	0.016	1.190	0.015	1.225	0.036	1.198	0.009	1.206	0.016
Income (cat. 4)	1.292	0.007	1.134	0.017	1.264	0.016	1.144	0.038	1.279	0.010	1.151	0.017
Income (cat. 5)	1.205	0.008	0.918	0.018	1.177	0.017	0.872	0.043	1.199	0.011	0.945	0.019
Married/Cohabiting	1.201	0.004	0.987	0.009	1.227	0.009	1.023	0.022	1.195	0.006	0.989	0.010
Immigrant	0.651	0.008	0.991	0.017	0.628	0.019	0.979	0.042	0.652	0.012	0.993	0.018
UI history (cat. 2)	1.160	0.005	0.978	0.013	1.186	0.012	0.893	0.030	1.105	0.008	0.983	0.014
UI history (cat. 3)	1.075	0.006	0.943	0.014	1.124	0.013	0.812	0.035	0.999	0.009	0.946	0.015
UI history (cat. 4)	1.032	0.007	0.830	0.016	1.178	0.017	0.684	0.043	0.958	0.010	0.821	0.016
Sickness/disability history	0.558	0.007	1.216	0.016	0.552	0.018	1.258	0.035	0.527	0.011	1.166	0.017
UI eligibility ¿ 1 year	0.806	0.006	1.342	0.014	0.826	0.013	1.313	0.032	0.808	0.008	1.330	0.015
UI eligibility ¿ 2 year	0.643	0.008	1.465	0.017	0.629	0.018	1.474	0.040	0.656	0.011	1.436	0.018
Unemployed hours	1.248	0.010	1.283	0.024	1.392	0.024	1.513	0.058	1.218	0.015	1.276	0.025
Region 2	0.971	0.007	0.758	0.015	0.941	0.014	0.842	0.032	0.987	0.010	0.769	0.016
Region 3	1.026	0.007	0.617	0.016	0.970	0.014	0.577	0.035	1.012	0.010	0.642	0.016
Region 4	0.976	0.007	0.746	0.017	0.904	0.016	0.561	0.039	1.016	0.011	0.816	0.018
Region 5	1.012	0.007	0.504	0.017	0.944	0.015	0.406	0.039	1.074	0.010	0.529	0.017
Region 6	0.932	0.007	0.517	0.018	0.881	0.015	0.469	0.039	0.962	0.011	0.538	0.018
Duration dependence:												
Months 1-3	1		1		1		1		1		1	
Months 4-6	1.528	0.005	1.460	0.010	1.218	0.013	1.142	0.032	1.663	0.006	1.445	0.010
Months 7-9	1.141	0.006	1.997	0.015	0.842	0.022	0.850	0.146	1.243	0.010	1.899	0.014
Months 10-12	0.902	0.008	2.474	0.025	0.701	0.034			0.941	0.015	2.223	0.022
Months 13-18	0.730	0.009	2.704	0.043	0.359	0.051			0.760	0.021	2.030	0.041
Months 19-24	0.609	0.014	2.646	0.089	0.285	0.072			0.633	0.062		
Months 25-	0.519	0.020			0.212	0.105						
Unobserved heterogeneity												
v_1	0.0031	0.024	0.0000	0.878	0.0007	0.046	0.0001	2.571	0.0030	0.030	0.0000	3.255
v_2	0.0024	0.025	0.0006	0.199	0.0006	0.080	0.0014	0.261	0.0022	0.031	0.0004	0.056
p_1	0.534				0.390				0.616			
p_2	0.466				0.610				0.384			
Observations	582,580				112,678				428,160			

Both equations also contain 8 industry dummies and quarter-of-year fixed effects.