The Exhaustion of Unemployment Benefits in Belgium. Does it Enhance the Probability of Employment? [‡]

Bart Cockx[§] and Jean Ries[¶]

June 1, 2004

Abstract

In Belgium unemployment insurance benefits can only exhaust for one category of workers: partners of workers with (replacement) labour income (mostly women) may loose their entitlement after an unemployment duration ranging from two to eight years, depending on individual characteristics. We contrast three propensity score matching estimators of the impact of benefit exhaustion on the probability of employment : a standard, a before-after and an IV matching estimator. We conclude that benefit expiration is anticipated as from the moment at which the worker is notified, three months in advance, and that it gradually increases the employment rate up to 25 percentage points 14 months after benefit withdrawal.

Keywords: unemployment insurance, programme evaluation, before-after estimator, nonparametric methods

JEL classification: J65, J68, C14

[‡]The authors acknowledge financial support of the Federal Ministry of scientific research (contract SSTC n° SO/10/036) and the National Employment Office (ONEM/RVA) for delivery of the data. Bart Cockx is grateful for the hospitality of the Tinbergen Institute in Amsterdam and CESifo in Munich where part of this research was realised. The authors thank Muriel Dejemeppe, Christian Göbel, Mareva Sabatier, Bas Van der Klaauw and Bruno Van der Linden for their comments. The usual disclaimer applies.

[§]Institut de Recherches économiques et sociales (IRES) and Department of Economics, Université catholique de Louvain, IZA, Bonn and CESifo, Munich. *Adress:* Place Montesquieu 3, B-1348 Louvain-la-Neuve, Belgium. Tel.: +32 10 473439; fax: +32 10 473945. E-mail: cockx@ires.ucl.ac.be.

[¶]Institut de Recherches économiques et sociales (IRES) and Department of Economics, Université catholique de Louvain. *Adress:* Place Montesquieu 3, B-1348 Louvain-la-Neuve, Belgium. Tel.: +32 10 473071; fax: +32 10 473945. E-mail: ries@ires.ucl.ac.be.

1 Introduction

The Belgian Unemployment Insurance (UI) system is unique in the world in that the receipt of unemployment benefits (UB) is not limited in time. There is, however, an exception to this rule. For partners of workers with (replacement) labour income (mostly women), benefits may under some conditions expire beyond some threshold unemployment duration. This paper aims at estimating the average treatment effect of benefit exhaustion on the probability of employment for women belonging to this group.[1] For this purpose we use an existing database that was constructed for evaluation purposes (see De Lathouwer *et al.* (2000)). It consists of two samples. One is drawn from the flow of UB recipients for whom benefits have expired between March and June 1997. The second sample is drawn from the stock of long-term unemployed workers in March 1997 for whom the benefits were not yet withdrawn. The database combines survey and administrative information. It allows to construct the employment history of these workers from 63 months prior to the sample selection date to 14 months after. On the basis of simple logit estimations, De Lathouwer *et al.* (2003) find that the scheme boosts employment rates by 14 percentage points 3 months after benefit expiration, decreasing to 9 percentage points after 15 months. These effects are significantly different from zero at the 95% level of significance. Our paper verifies whether these results uphold if one uses evaluation methods that require less restrictive identifying assumptions.

Job search theory predicts that benefit exhaustion generates three effects (see Mortensen (1977, 1990), van den Berg (1990), Fredriksson and Holmlund (2003b) for a survey). The two first effects realise ex post, once the entitlement has expired. First, as the worker has an interest in maintaining her living standard, the absence of unemployment benefits enhances incentives to search for and to accept jobs. Second, since the worker is no longer eligible, she has an additional interest in being hired and in remaining employed until she can re-qualify for UB [2]. Thirdly, since the moment at which one finds a job is not deterministic [3], workers will alter their behaviour well in advance of the UB exhaustion. To the extent that workers are aware of future benefit exhaustion, they anticipate this event by starting to search for a job beforehand as to preclude the income loss in the event that a job is not timely found. Theory predicts that this anticipation will gradually increase the employment hazard as one approaches the expiration date. However, if initially workers are unaware of the finite entitlement length, the employment hazard will jump upwards from the point at which they become informed and subsequently increase gradually as for the informed workers. Due to the lack of transparency of the Belgian scheme, we argue below that the latter case applies: the behaviour will only change as from the moment at which the (unemployed) worker is notified by the Unemployment Agency (UA). This occurs roughly three months before benefit exhaustion.

[1] Men were not retained, since they form a too small subsample of treated workers to perform any sensible statistical analysis.

[2] Note that this effect disappears once the worker has completed the qualifying period.

[3] Note that this may not be true if the worker is recalled by his previous employer, see e.g. Katz and Meyer (1990).

There exists a vast empirical literature that tries to test these theoretical predictions. Devine and Kiefer (1991), Atkinson and Micklewright (1991), Layard *et al.* (1991), and more recently, Holmlund (1998), Pedersen and Westergård–Nielsen (1998) and Fredriksson and Holmlund (2003a) survey this literature. The first generation of studies, realised both in the US and in Europe between the seventies and the early 1990s, generally confirm the predictions of the theoretical model, but impacts are generally quite modest. Layard *et al.* (1991) state that "the basic result of these studies is that the elasticity of expected duration with respect to benefits is generally relatively low, situated in a range between 0.2 and 0.9". The problem with this earlier literature is that, due to methodological flaws, the estimations cannot always be interpreted as the true causal effect of a variation in the unemployment compensation.[4] The last ten years the evaluation literature has evolved enormously and generally performs much more careful analysis in this respect. Remarkably, the more recent (European) studies (Dormont *et al.* (2001) for France, Carling *et al.* (2001) for Sweden and Roed and Zhang (2003) for Norway)[5] seem to find a much larger impact of UB on the unemployment duration with elasticities ranging between 0.95 and 1.6.

Another strand of the empirical literature studies the impact of benefit exhaustion on the profile of the hazard rate to employment as well as the impact of variation in the length of entitlement period. These studies tend to support the last mentioned findings in that impacts are generally found to be important and consistent with job search theory. Almost every study [6] reports increases of the job finding rate as benefit exhaustion is approached (Wurzel (1990); Lindeboom and Theeuwes (1993); Carling *et al.* (1996); Joutard and Ruggiero (1996); Thoursie (1998); Bratberg and Vaage (2000); Dormont *et al.* (2001); Roed and Zhang (2003)). Moreover, all studies find that the extension of the entitlement duration significantly reduces the rate at which unemployment is left for employment (Moffit (1985); Ham and Rea (1987); Katz and Meyer (1990); Hunt (1995); Winter-Ebmer (1998); Card and Levine (2000); Lalive and Zweimüller (2002)).

A final piece of evidence, suggesting that monetary incentives matter, comes from the sparse evidence of the impact of UB sanctions on the employment hasard. Sanctions are punitive benefit reductions that are used to enforce compliance of UB claimants to job search requirements. An important feature of a sanction is the benefit reduction itself. In that sense effects of sanctions indirectly provide evidence for the effects of benefit reductions. However, one should be careful with this interpretation : "the monitoring and sanctioning regime itself can be expected to affect individual behavior, and the incentives to comply with the rules may increase as the regime is typically tightened after a violation of the rules" (Abbring *et al.*, 2000, p.3). Nevertheless, the benefit cut undoubtedly represents a significant component of the impact of a sanction and can be considered as a useful complementary evidence. Abbring *et al.* (2000) find very large impacts of sanctions in the Unemployment Insurance Scheme in the Netherlands. The

[4] See also Abbring *et al.* (2000, p.26–27) on this point.

[5] Neither Jenkins and Garcia-Serrano (2000) or Bover *et al.* (2002) seem to find large elasticities for Spain, however.

[6] Again the study of Jenkins and Garcia-Serrano (2000) is an exception in that the employment hazard in Spain does virtually not change as one approaches benefit expiration.

implied absolute values of the elasticities of the employment hazards with respect to the benefit levels would range between 2 and 5. Van den berg et al. (2004) find similar results for sanctions on transitions from welfare to work. Finally, Lalive *et al.* (2002) and Jensen *et al.* (2003) find also quantitatively important effects of benefit sanctions in respectively the Swiss and Danish UB scheme.

Our paper contributes to this literature by providing a new estimate of the impact of benefit exhaustion on the basis of modern evaluation methods (See e.g. Heckman *et al.* (1999), Blundell and Costa-Dias (2002)). As to enhance the credibility of our findings, we estimate the impact on the basis of three different methods: a standard, a before-after and an Instrumental Variables (IV) propensity score matching estimator. Since all three methods are based on matching, they are not sensitive to functional form or parametric assumptions and they take care of individual heterogeneity in the treatment effect of the treated.

The paper is structured as follows. In Section 2 we briefly describe the system of benefit exhaustion in the Belgian UI scheme. In Section 3, we describe the database used in our study. Section 4 is explains the estimation methods and results. A last section concludes.

2 Benefit Exhaustion in Belgium

As in most European countries, in Belgium workers qualify for UI after involuntary redundancy if their record of salaried employment is sufficiently long.[7] However, unlike in most countries, once workers qualify for UI in Belgium, if they are "available for the labour market" and if they comply to certain administrative rules, they are entitled to benefits for an indefinite period. There is only one exception to this rule: benefits may be withdrawn after a an unemployment duration, ranging from two to eight years, if one is less than 50 years old and partner of someone with a (replacement) income exceeding a particular level.[8] The precise value unemployment duration at which benefits exhaust depends on the age class (3), the region of residence (30) and the gender of the unemployed person[9]. The thresholds are calculated as 1.5 times the average unemployment duration for each class *, as measured at a certain point in the past.*[10] From Table 1, one can deduce that these durations are quite heterogenous across and within age classes.

For the subsequent analysis it is crucial to realise that unemployment duration is not measured according to the standard "ILO" or "EUROSTAT" definitions. The counter accumulates any day of unemployment since the first entry, days of part-time (un)employment being weighted appropriately. It is only reset to zero

[7] One particularity of the Belgian system is that school-leavers are also entitled, after some waiting period.

[8] For this category of workers the unemployment benefit is degressive, decreasing in three phases: it starts at a replacement rate of 55% during the first 12 months (with a floor and a ceiling of respectively 540 EUR and 920 EUR in January 2004), falling to 40% (within the range of 540-670 EUR) during the next three months (possibly longer according to the employment experience) and finally falls to a flat rate level (380 EUR in January 2004). Benefit exhaustion always occurs in the last phase.

[9] Thus, there are $3 * 30 * 2 = 180$ different thresholds.

[10] At which point in time these unemployment durations were calculated is unclear. It is clear, however, that they haven't been updated for a long time.

if there is a period of at least two years of full-time employment between two subsequent unemployment spells. It is not affected by temporary interruptions of unemployment.

Table 1: Threshold Durations for Women (in months)

Age	*< 36*	*36–45*	*45–50*
Mean	50	60	79
St.Dev	10	11	13
Min.	30	35	45
Max	74	84	99

An unemployed worker is notified not later than three months before the potential benefit exhaustion, i.e. the date at which the unemployment duration crosses the threshold. From that moment onwards, one may appeal against that decision on the following grounds. (1) If the annual net taxable family income (excluding the benefit) is below a certain threshold. In January 2003, this threshold was 16,750 EUR, i.e. roughly 140% of the minimum wage, augmented by 670 EUR, for each dependant person of the household. (2) If the unemployed worker can prove "an exceptional and continuous job search effort throughout the unemployment spell". (3) If the unemployed worker is participating in an active labour market policy. (4) If the unemployed worker can prove a substantial degree of disability. If the appeal is justified, benefit exhaustion is postponed definitely or temporarily, for a defined or undefined time period. As consequence, between 1991 and 2003 only roughly 40% of those notified had their benefits effectively withdrawn. In 1997, this concerned slightly less than 19,000 individuals or roughly 4% of the total number of benefit recipients during that year.

Once benefits are withdrawn, the worker can re-qualify in one of the three following cases : (1) the worker separates from his/her partner; (2) the yearly net taxable income drops below a lower threshold (in January 2003 14,795 EUR plus 670 EUR per additional dependant person); (3) proof of 312 days of full time employment over an uninterrupted period of 18 months is delivered. As opposed to other countries, there is no unemployment assistance scheme that sets in when benefits exhaust. In principle, the individuals can claim means-tested welfare assistance. However, since only workers with family income above some threshold can loose UI benefits, this only very rarely occurs [11]: within the sample analysed in this paper only 0.3% claimed welfare benefits within the first 4 months following the expiration (De Lathouwer *et al.*, 2003, p.91).

From this description it is clear that the rules pertaining to benefit exhaustion are complex. Indeed, the precise moment of benefit exhaustion depends on a number of criteria that may vary over time: partner-status, place of residence, family income, age, etc. Moreover, for workers with irregular employment histories it is rather difficult to keep track of their effective elapsed unemployment duration. We therefore claim that it is unlikely that workers anticipate the benefit exhaustion

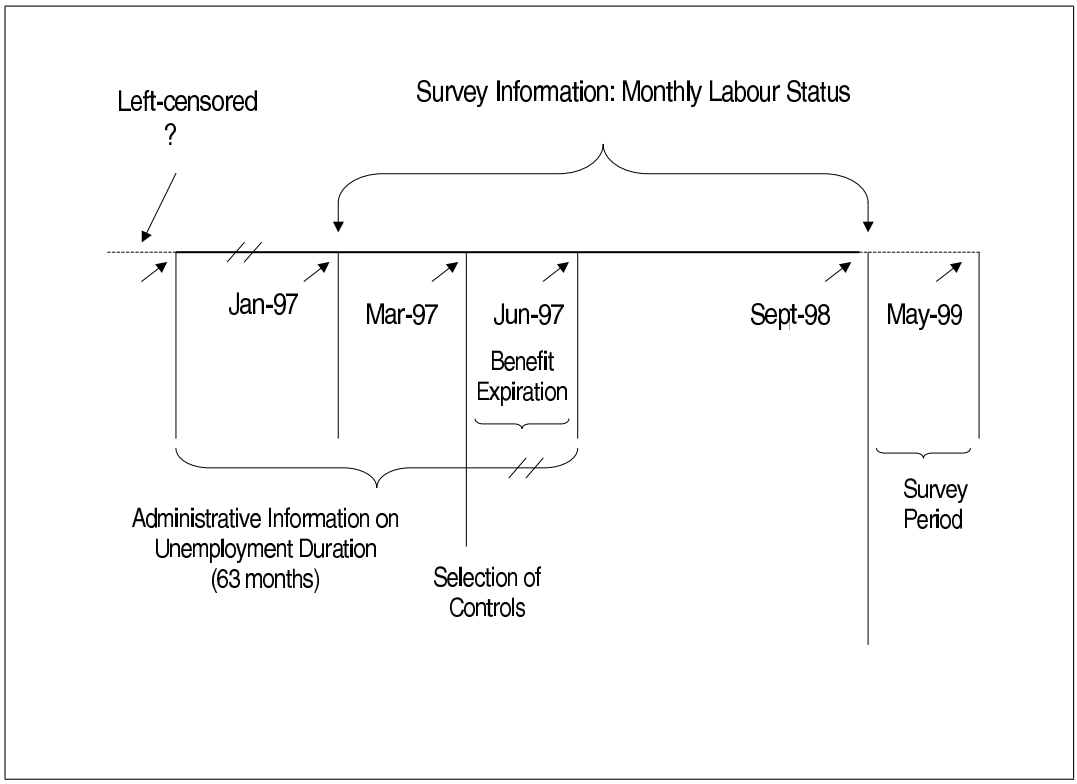
[11] e.g. if there is a recent drop in family income that is not captured by the past annual net taxable income.

before being notified. This claim cannot be rejected on the basis of the empirical analysis below.

3 Data

The analysis is based on an existing database that was constructed for the purpose of evaluating the benefit exhaustion scheme (see De Lathouwer *et al.* (2000)). The structure of this database is rather complex. So as to facilitate the explanation, we summarised its features in Figure 1. The database consists of two samples of unemployed workers. Since in both samples the workers are less than 50 years old and cohabiting with a partner earning some (replacement) labour income, benefits can potentially expire for both groups. The first sample, the *treatment group*, contains 1,506 individuals and is randomly drawn from the flow of UI recipients for whom benefits effectively expired between March and June 1997. The second sample, *the control group*, contains 1,205 individuals and is drawn from the stock of "long-term" [12] unemployed workers in March 1997 from whom the benefits were not yet withdrawn.

Figure 1: The Time Structure of the Database



[12] "Long-term" means here that the worker has experienced unemployment - possibly with interruptions - during more than 18 months since January 1992.

We imposed three additional sample selection criteria. First, even if the response rate was fairly high (66% for the treatment group and 73% for the control group), it was not complete. In the analysis we do not account for the potential selection bias on unobserved characteristics induced by non-response. There is, however, no reason for serious concern, since no significant selection could be detected on the basis of the *observed* individual characteristics (see De Lathouwer *et al.* (2000) for details). Second, since only few partners of unemployed men earn a (replacement) labour income, men are a too small sub-sample (169 among the treated respondents and 196 among the controls) to perform a sensible statistical analysis. Only women are therefore retained. Thirdly, we excluded 188 "polluted controls" (=27% of the female respondents in the control group). These refer to workers whose benefits were not yet suspended at the sampling date in March 1997, but for whom one of the three following events occurred between the sampling and the survey date: (1) they were notified, but benefits did not (yet) exhaust; (2) they were notified and benefits expired effectively; (3) benefits were withdrawn for another reason (e.g.: refusing a job offer, misrepresentation of household composition as to obtain higher benefits,...). The survey allows to identify these individuals, but, unfortunately, does not inform on the date at which the above-mentioned events occurred. After imposing these selection criteria, we retain 826 women in the treatment group and 404 in the control group.

Both samples were drawn from administrative files on UI recipients. These contain information on individual characteristics (the age, an indicator for the Belgian nationality, the highest level of education attained and the sub-region of residence) and on unemployment histories during up to 63 months prior to the sample selection date. The administrative information was completed by responses to a survey conducted at one point in time between September 1998 and May 1999.[13] The reported number of children in the household at the survey date was the only additional individual characteristic retained. Besides, the survey respondents reported their monthly labour market status between January 1997 and the date of the survey.

The data contain three major deficiencies for our purposes. First, there is a break in the series regarding the labour market histories, since it comes from two sources: prior to the sampling date from the administrative source and afterwards from the survey. Second, they do not contain a variable that defines unemployment duration in exactly the same way as the administrative rules regarding benefit expiration. Thirdly, the data do not inform at which moment unemployed women are notified of the benefit exhaustion.

As to the last deficiency, we only know that the rules impose that the unemployed workers must be notified not later than 3 months before benefits expire. In the sequel we simply assume that it occurs at that moment. The first two deficiencies are addressed differently according to the statistical method applied to the data. The benchmark "before-after" method described in the next section only analyzes the treatment sample. For this group, we assume that the effective elapsed duration at the moment of expiration coincides with the theoretical thresh-

[13] The large discrepancy between the starting and the ending date of the survey was a consequence of some budgetary problems as explained in De Lathouwer *et al.* (2000).

old duration at which benefits should be withdrawn according to the rules. The elapsed duration prior to benefit exhaustion is then calculated backwards starting from the date of benefit withdrawal on the basis of the 63 monthly administrative information regarding the number of days of benefit receipt. After this date one month of unemployment duration is added for each month the female workers reported to be unemployed or inactive according to the monthly labour market status of the survey's questionnaire.

In the causal analysis below the monthly employment status [14] is the outcome variable. After January 1997, the monthly employment status corresponds to the one reported by survey respondents at the time of the interview. However, the "before-after" estimator requires information prior to this date. We therefore construct a series predicted on the basis of an estimated relationship between the monthly employment status reported in the survey and the number of days unemployed every month according to the administrative files. Such an estimation is possible, since there is a period of overlap between January 1997 and June 1997, during which information of both sources is available. Details of the estimation can be found in Appendix B. The fit during the period of overlap is close.

For the sensitivity analysis to the benchmark "before-after" estimator, the above-mentioned data problems are resolved differently. First, we estimate the impact of benefit exhaustion only as from the moment at which benefits are withdrawn and not before. As a consequence, the outcome variable can be exclusively based on the employment histories reported in the survey. Second, we cannot follow the same procedure as to define the elapsed unemployment duration, since for the control sample the date of effective benefit exhaustion is unknown. We could estimate this duration on the basis on the administrative unemployment record in the 63 months prior to the sampling date in March 1997, i.e. as from January 1992. However, we face a problem of left censoring. By lacking information on the unemployment history prior to 1992, we tend to underestimate the unemployment duration determining the timing of benefit exhaustion. This problem is not negligible: for a sizeable proportion of the treatment group (67%) the elapsed duration as calculated from the administrative files is lower than the threshold duration at which benefits should exhaust according to the rules. In order to resolve this problem we related, in a regression analysis on the treatment group, the available explanatory variables to the extent to which unemployment duration, as measured by the available administrative files, is under-estimated. On the basis of this estimation, we then predict the extent to which the unemployment duration of the control group is under-estimated. Details of this correction can be found in Appendix A. Note finally that the corrected unemployment duration is not used in the analysis as such, but that it is only used to construct for each month of analysis an indicator variable that determines whether or not the benefits of some control unit could have be withdrawn according to the rules. As explained below, this indicator defines an instrumental variable (IV).

Table 2 reports descriptive statistics of the retained control and treatment sample. It turns out that the treated individuals are on average younger than the controls, slightly better educated, and have on average more dependant children in

[14] We do not distinguish between full- and part-time.

their households. As far as the labour market histories are concerned, the treated have a higher cumulated unemployment duration than the controls, and their last unemployment spell is substantially longer. The average local unemployment rate in the sub-region of residence is very close for both groups.

4 The Impact of Benefit Exhaustion on the Employment Rate

This paper aims at providing an estimate of benefit exhaustion on the women retained in the above-mentioned treatment sample. As to enhance the credibility of our findings, we estimate the impact on the basis of three different methods and contrast the results. A common thread in these methods is that they will all be based on propensity score matching methods. As such our estimation methods do not require functional form or parametric assumptions and take care of individual heterogeneity in the treatment effect of the treated.

The standard cross section matching method requires an assumption of unconfoundedness or "selection on observables" (see Rosenbaum and Rubin (1983); Heckman *et al.* (1997, 1998); Dehejia and Wahba (1999); Imbens (2003) for a recent review). Since in many applications, as for the one presented in this paper, this assumption is overly restrictive, one calls for the use of difference-in-differences (DD) matching estimators to eliminate any remaining fixed unobserved bias term (Heckman *et al.* (1997) ; Smith and Todd (2004)). Nevertheless, we do not implement a DD estimator in the present application, but rather a simple before-after Estimator, be it on an unusual time-scale. We justify this choice below.

One may be suspicious of the validity of the before-after estimator since it requires the strong identifying assumption that all systematic time-varying effects can be captured by a number of observed time-varying variables. In a sensitivity analysis we therefore contrast the findings of the benchmark before-after estimator to two alternative estimators : a standard matching and an IV estimator. Note, since we have a discrete outcome variable, Two Stage Least Squares will in general yield inconsistent estimates (see e.g. Wooldridge (2002, p.478)). We avoid this problem by implementing the IV estimator as a matching estimator (Heckman *et al.* (1999); Ichimura and Taber (2001)).

4.1 The benchmark model : the before–after estimator

4.1.1 A preliminary analysis

Figure 2 displays the evolution of the employment rate of the women in the treatment group. Calendar time is normalised to zero in the month that benefits exhaust, i.e. between March and June 1997. The time scale runs from 62 calendar months before the expiration date to 14 months afterwards.[15] The employment rate evolves according to a U-shaped pattern. Starting at a level of 24% at -62 it

[15]The survey response period starts in September 1998. This is 14 months after expiration for those women whose benefits were withdrawn in June 1997. Beyond 14 months we therefore no longer observe the employment status of all women.

Table 2: Sample Means of Explanatory Variables

	Controls (N = 404)	Treated (N = 826)
Individual Characteristics		
Age	35.67	32.81
Number of children in the household	1.22	1.73
Non-Belgian nationality	0.12	0.10
No diploma[a]	0.38	0.24
Lower secondary[a]	0.27	0.31
Higher secondary[a]	0.29	0.33
College[a]	0.06	0.08
Labour Market History in March 1997		
Unemployment duration[b]	45.74	53.69
Number of unemployment spells[c]	2.62	3.11
Duration of the last spell[d]	29.92	26.36
Last spell left censored[e]	0.27	0.19
Local Unemployment Rate[f]	0.22	0.23

All these variables come from administrative registers, with the exception of the "number of children in the household", which was reported by the individuals in the survey.

[a] "No diploma": less than 9 years of education; "Lower secondary": at least 9 years of education; "Higher secondary": at least 12 years of education; "College": at least 15 years of education

[b] Cumulative number of months passed in unemployment in the period from January 1992 to March 1997.

[c] Number of uninterrupted unemployment spells in the period from January 1992 to March 1997.

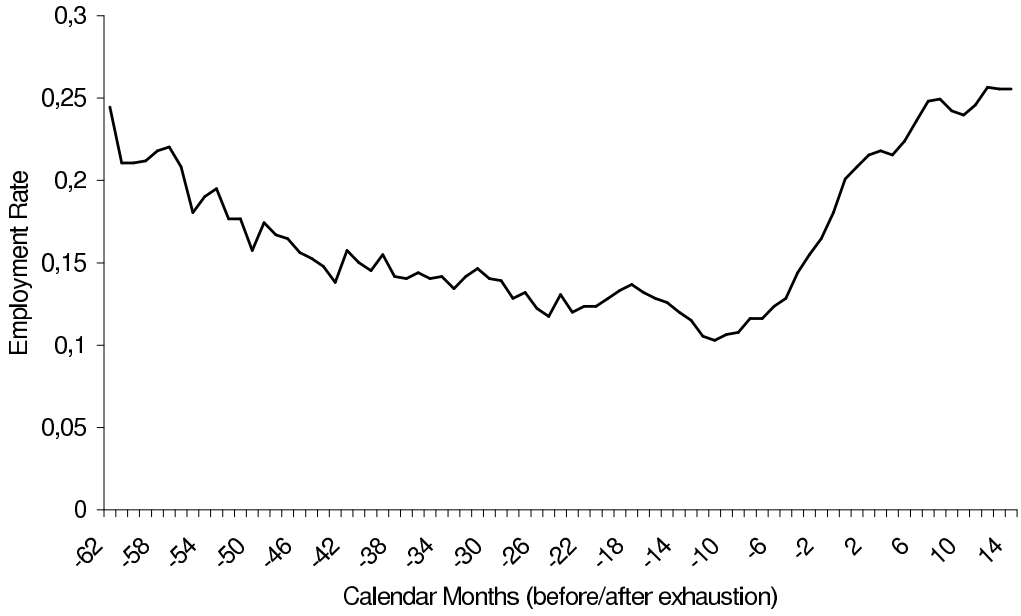
[d] Duration of the last unemployment spell, ending or still ongoing in March 1997.

[e] Binary Variable, indicating whether the last unemployment spell started prior to January 1992.

[f] Sub-regional unemployment rate for women as reported by the National Administration of UI (ONEM/RVA).

falls steadily to 11% at -11 months. Thereafter it increases continuously up to a level of 26% at the end of the observation period.

Figure 2: The Evolution of the Employment Rate over Calendar Time before and after Benefit Expiration



One might be tempted to interpret this graph as proof for the presence of a temporary "Ashenfelter-dip" in the employment rate (Ashenfelter, 1978). Alternatively, one may want to interpret the initial decrease as "permanent", induced by an unemployment dependency effect and the reversal of this evolution at -11 as an anticipated behavioural impact of the benefit exhaustion that sets in well in advance of the moment at which women are notified, three months beforehand. Neither of these interpretations is, however, correct. The observed pattern is the consequence of an incorrect time-scale prior to the expiration date.

Workers are notified three months of *benefit entitlement* prior to the expiration date. This corresponds to three months of *calendar* time only if workers don't leave unemployment during these last months of entitlement. However, in general workers may interrupt their unemployment spell temporarily. As already stated (see Section 2), this interruption only stops the counter of unemployment duration, but does not set it to zero. Consequently, one month of UI entitlement may cover a much longer calendar time period.

In Figure 3, the evolution of the employment rate is plotted as a function of *entitlement* months prior to the expiration month and of *calendar months* afterwards. With this new time scale the employment rate is no longer clearly defined, since one month of entitlement may span several calendar months. We propose two alternative definitions of the employment status of a treated woman:

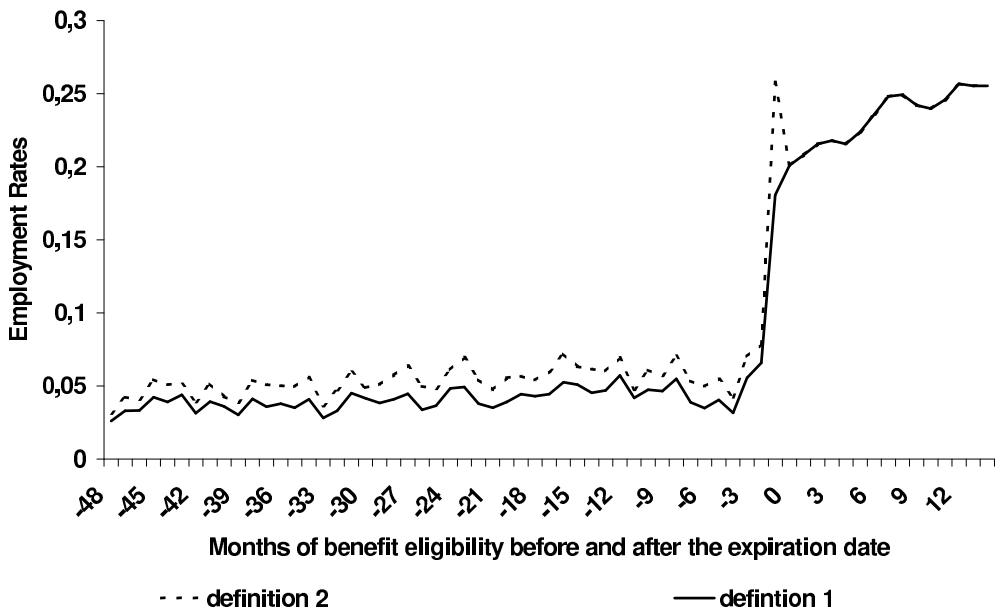
Definition 1 The employment status of a treated woman is defined as the *average* employment rate over the calendar months during which the benefit entitlement

duration remains constant.[16]

Definition 2 The employment status of a treated woman is equal to one if the woman has been reported employed during at least one month of the period during which the benefit entitlement duration remains constant.

The plotted employment rates in Figure 3 are averages of these employment indicators over the treatment sample. The employment rate according to the second definition is, by construction, always larger than the employment rate according to the first. Note that with this new time scale we cannot go as far back in time, since each month of benefit eligibility may correspond to several calendar months. In fact the sample is incomplete prior to 4 entitlement months before benefit exhaustion. This means that at least one woman has not been paid benefits during at least 63 months within the 4 last months that she was entitled to UI. The employment rate prior to this date is therefore based on an incomplete sample gradually decreasing in size. [17]

Figure 3: The Evolution of the Employment Rate over Eligibility Duration before Benefit Expiration and of Calendar Time Afterwards



The figure reveals that the evolution over this new time-scale of the employment rate (according to both definitions) is completely different. The employment rate is now roughly constant until -4. Three months prior to benefit expiration, i.e. roughly at the moment at which the unemployed woman is notified, the employment rate suddenly starts to increase. It rises most steeply one month prior to the month in which benefits are withdrawn. On the basis of the second definition the employment rate is even higher in this month than in the following

[16]The entitlement duration is rounded off up to the nearest month.

[17]This sample size is reported up to -13 in Table 3 in the following sub-section. Complete information is available upon request.

one. This can be explained as follows. Unemployment benefits of all sampled women will eventually exhaust by construction, as a consequence of the sampling scheme. All employment prior to the expiration date must therefore be temporary and the employment rate is bound to decrease in the month in which benefits are withdrawn.

Figure 3 provides unambiguous evidence of a positive causal impact of benefit exhaustion on the employment rate: the increase of the employment rate one month prior to the exhaustion is too abrupt to be credibly explained by other time-varying factors, such as labour market conditions, unemployment duration or life-cycle events. To estimate the quantitative magnitude of this impact we nevertheless need to purge the gross impact of these time-varying factors, in particular if we wish to estimate the long-term impact, for instance 14 months later. We propose two purging methods in the next subsection.

Figure 3 also proves that the benefit exhaustion is anticipated. As a consequence, the impact as estimated on the available sample is necessarily a lower bound: women who manage to escape benefit exhaustion, for instance by timely finding a full-time job lasting more than two years, are not retained in the treatment sample. We need other data to gauge the importance of this anticipation effect.

4.1.2 Methods to Control for Time-Varying Variables

The analysis of the preceding section suggests that, if applied on the modified time-scale, two necessary identifying conditions of a before-after estimator are satisfied in this empirical application. First, the beginning of the treatment period is clearly identified at three months of benefit entitlement before the expiration date. Second, on the modified time-scale, the outcome variable does not display any longer a temporary "Ashenfelter dip" prior to the start of the treatment.

Nevertheless, a before-after estimator is only an unbiased estimator of the impact to the extent that it purges the outcome variable of other time-varying determinants. The DD estimator is a popular solution to this problem. We do not apply this method in our application for several reasons. First, the DD estimator requires that the bias between control and treatment group remains constant (Eichler and Lechner, 2002). However, it is not clear how this constancy can be guaranteed in a "nonlinear environment": even if the bias is induced by fixed determinants, the non-linear transformation of this fixed term evaluated at two different time points is generally no longer fixed. Since the outcome variable is (partly) discrete in our empirical application, this critique applies.[18] Blundell *et al.* (2001) solve this problem by assuming that a nonlinear transformation (e.g. the inverse logistic) of the expected outcome variable can eliminate the bias. However, it is not clear which is the appropriate choice of this transformation function and how to generalise this to more than two outcomes (Athey and Imbens, 2002, p.24).

There is an additional reason why we do not implement DD. On the basis of the modified time-scale a DD estimator is no longer feasible. The reason is

[18] In the empirical analysis the outcome variable is the employment status as defined according to definition 1 in the previous section. Since this variable is an average, it is only *partly* discrete.

twofold. First, the data on control units only contain imperfect information on the entitlement duration left until expiration (see Section 3). Second, even if we had this information, these units would no longer be controls, since if they are measured at the same moment as the treated according to this new time scale, their benefits would exhaust at the same moment as for the treated.[19]

We therefore control for time-varying variables in a different way. We assume that there are essentially two important time-varying determinants of the employment rate: the state of the local labour market and the elapsed cumulative unemployment duration.[20] We believe that other factors, such as those pertaining to the life-cycle, affect the population randomly at the different time periods, so that they average out.

One may question whether it is appropriate to control for the unemployment duration at the time of treatment. For, the observed unemployment duration after treatment is lower than the counterfactual unemployment duration in the absence of treatment: if benefit exhaustion enhances the time spent in employment, then the cumulative unemployment duration increases less than if benefits did not expire. In principle, we should therefore condition in the after treatment period on this higher counterfactual duration, but we cannot, since it is unobserved. By this neglect we *under-estimate* the impact of benefit exhaustion on the employment rate. As this indirect impact grows over time, this bias will increase with k. However, this bias cannot be large. To obtain some sense for the order of magnitude we estimated the treatment effect excluding the unemployment duration as a control. This reduces the impact estimate at $k = 14$ by about 1.5 percentage points.

We control for these time-varying variables in two ways. The first method is described in Heckman *et al.* (1999, p.1985) in which the outcome variable is adjusted by some function of the time-varying variables. In principle, one could estimate this function non–parametrically, but this may not be easy, especially if the outcome variable is discrete. We therefore propose an alternative non-parametric procedure. This adjusts the outcome variable at each moment by a method of matching on the propensity score of the time-varying variables. A drawback of the method is that a common support problem is bound to occur. We solve this problem by implementing the method as a sum of "incremental" matching estimators applied to subperiods of the period under consideration.

For a formal analysis, we introduce some notation. First note that the variables below are all defined for women belonging to the treatment sample. To avoid burdensome notation, we ignore everywhere the reference to these individuals, usually denoted by a subscript i.

Let k denote time according to the modified scale: if k is negative it means that $|k|$ months of entitlement are left until benefits exhaust; a positive k refers to the number of calendar months after the expiration. D is an indicator equal to one

[19] This could be avoided if one constructed a hypothetical threshold duration for a group of women who are never eligible to the treatment, e.g. singles or women cohabiting with a partner without (replacement) income of employment. However, these data were not available.

[20] The first is captured by the sub-regional unemployment rate for women as reported by the National Administration of UI (ONEM/RVA). The second is captured by the cumulative (possibly interrupted) unemployment duration since April 1992. These variables are evaluated at the first calendar month of a period in which the duration of remaining benefit eligibility remains constant.

if the individual is "treated" and to zero "otherwise". "Treated" means that one is informed of the remaining entitlement duration or that benefits have effectively been withdrawn. We assume that treatment starts when the unemployed worker is notified, i.e. at $k = k^*$, and that $k^* = -3$. The latter assumption can be tested, as shown below.

Y_{1k} and Y_{0k} denote the potential outcomes of respectively being treated or not at time k. The observed outcome Y_k is then defined as:

$$Y_k \equiv DY_{1k} + (1 - D)Y_{0k} \tag{1}$$

We aim at estimating the Average Treatment Effect on the Treated at time k:

$$ATT_k \equiv E(Y_{1k} - Y_{0k}|D = 1) \tag{2}$$

A Method based on an Adjustment Function

Consider a deterministic function $g(x_k)$ that adjusts for the realisations x_k of a random vector X_k of time-varying variables evaluated at time k. This function can be estimated on the basis of the pre-treatment ($k < k^*$) observations of the treatment sample. The adjusted potential outcome then takes the following form:

$$A_{pk} \equiv Y_{pk} - g(x_k) \tag{3}$$

where $p = 0, 1$. To identify ATT_k , two assumptions must be satisfied:

Assumption 1 $\forall k' < k^* : E(A_{0k'}|D = 1) = E(A_{0k'}|D = 0)$

This assumption means that in the pre-treatment period the treatment indicator may not systematically affect the adjusted potential no-treatment outcome. As such, we can replace the average potential adjusted outcome of the treated $E(A_{0k'}|D = 1)$ by the average observed adjusted outcome of the same women prior to treatment $E(A_{k'}|D = 0)$. Since the expectations are taken over the same group of individuals, the assumption can only be violated if there is an unobserved determinant of the outcome variable that systematically changes simultaneously with the treatment status. This is very unlikely to occur in the current empirical application.

Assumption 2 $\forall k, k' : E(A_{0k}|D = 1) = E(A_{0k'}|D = 1)$

This assumption implies that the adjustment function $g(X_k)$ captures all systematic time-varying factors affecting the employment status of the treated women. Equivalently, the assumption implies that the distribution of unobserved factors that are not captured by this function does not systematically vary over time. [21]

Proposition 1 [22] *If Assumptions 1 and 2 hold then:*

$$\forall k' < k^* : E(A_{k'} - A_{k^*-1}|D = 0) = 0$$

[21] Note that, strictly speaking, Proposition 2 below only requires that there exists at least one k' in the pre-treatment period for every k in the treatment period for which the statement in Assumption 2 is valid. However, the more restrictive formulation is required for proof of Proposition 1.

[22] see Appendix C for a proof.

If Proposition 1 cannot be statistically rejected, this provides support for the hypothesis that the before-after estimator indeed identifies the causal impact of the treatment. It is no proof, however. For instance, the true adjustment function may change after the start of the treatment.[23]

The next proposition states that we can estimate the ATT_k by taking the difference between the average adjusted observed outcome in treatment month k and the average adjusted observed outcome at some moment k' in the pre-treatment period.

Proposition 2 [24] *If Assumptions 1and 2 hold then:*

$$\forall k \geq k^*, \exists k' < k^* : ATT_k = E(A_k|D = 1) - E(A_{k'}|D = 0)$$

A Method Based on Matching

The method just described requires the adjustment function to be correctly specified. This can be indirectly tested on the basis of Proposition 1. However, the power of this test is unknown. An alternative consists in estimating $g(.)$ non-parametrically, but this may be complicated if the number of regressors is high or if the outcome variable is discrete, as in the present empirical application. We propose an alternative non-parametric procedure based on a matching method. We propose two alternatives: "*direct*" and "*incremental*" matching.

Direct Matching

In order to identify the ATT_k according to this alternative method, we modify the above assumptions and we add a third one. First, we assume that the treatment status may not systematically affect the potential no-treatment outcome given some realisation the observed time-varying variables:

Assumption 3 $\forall k' < k^*, \forall x_{k'}$:

$$E(Y_{0k'}|D = 1, X_{k'} = x_{k'}) = E(Y_{0k'}|D = 0, X_{k'} = x_{k'})$$

Second, we assume that, conditional on a realisation of the observed time-varying variables, the potential no-treatment outcome for the treated does not systematically vary over time:

Assumption 4 $\forall k, k', x_k : E(Y_{0k}|D = 1, X_k = x_k) = E(Y_{0k'}|D = 1, X_{k'} = x_k)$

This assumption means that for every realisation of the time-varying variables, x_k, the distribution of unobserved determinants of the outcome variable is constant over time. This assumption is violated if x_k does not capture all time variation or if the composition of a sub-sample with a particular realisation of x_k changes over time. The first type of violation can be, for instance, induced by a policy change after a particular date. A violation of the second type can be the result of a movement from a an economic downturn to an economic upturn. If one conditions

[23] See Heckman and Hotz (1989) for discussion.

[24] see Appendix C for a proof.

on a high local unemployment rate, in a downturn the sub-sample of unemployed workers may contain a much larger share of qualified workers than in an upturn.[25] Note that in this respect Assumption 4 is more restrictive than 2. Finally, we must make a common support assumption (see e.g. Lechner (2001)):

Assumption 5 $\forall\ k\ ,\ \forall\ x_k\ ,\ \exists\ k' < k^*\ :\ Prob(X_{k'} = x_k) \neq 0$

This means that, for every realised values x_k of the vector of time-varying variables in treatment month k, there must in principle exist an identical realisation in month k' in the pre-treatment period.[26] This is required to find a control unit for each realised treatment. The larger the interval between the treatment and the pre-treatment period, the more likely this assumption is violated. In fact, we report below that for the present empirical application there is a substantial common support problem. As a consequence, we can not estimate the causal impact of benefit exhaustion of a large share (30%) of the treated women. This neglect could seriously bias the estimator ATT_k of the *total* treated population. This is why we will propose a variant of the method – "incremental matching" – that recovers the causal impact for the excluded treated units.

As for the previous method, two propositions summarise how the Assumptions 3 and 4 imply testable outcomes in the pre-treatment period on the one hand and how they identify ATT_k on the basis of a matching estimator on the other hand.

Proposition 3 [27] *If Assumptions 3-5 hold then:* $\forall\ k' < k^*, \forall\ x_{k^*-1}$:

$$E[E(Y_{k'}|D = 0, X_{k'} = x_{k^*-1}) - E[Y_{k^*-1}|D = 0, X_{k^*-1} = x_{k^*-1})|D = 0] = 0$$

Thus, conditional on the time-varying factors, the outcome variable may on average not vary in the pre-treatment period.

Proposition 4 [28] *If Assumptions 3-5 hold then:* $\forall\ k \geq k^*,\ \exists\ k' < k^*$:

$$ATT_k = E[E(Y_k|D = 1, X_k = x_k) - E(Y_{k'}|D = 0, X_{k'} = x_k)|D = 1]$$

where the outer expectation is over the distribution of $X_k|D = 1$

This proposition states that we can estimate the impact by matching each treated woman to a woman who has not yet been treated, but who has the same realisation of the time-varying variables. If these variables take on discrete values, one can match directly on each realisation of the time-varying variables. Alternatively, we can use the results of Rosenbaum and Rubin (1983) to justify matching on the propensity score.

Incremental Matching

[25] In the empirical application we do not only condition on the unemployment rate, but also on the unemployment duration. Since the latter variable is likely to be correlated with the level of qualification, this may mitigate the problem.

[26] The literature usually states this assumption as follows:
$Prob(D = 1|D = 0 \text{ or } D = 1, x_k) < 1$.

[27] see Appendix C for a proof.

[28] see Appendix C for a proof.

We mentioned that, as a consequence of lack of common support, the direct matching method may not be very appealing. This is why we propose a variant that eliminates the problem, be it at the cost of reducing the precision of the estimator. In a nutshell, incremental matching consists in taking the sum of the impacts of moving, *incrementally*, from one time period to the next. If we consider a sufficient number of increments, this sum measures the impact of a treatment at some moment substantially beyond the start of the treatment. The common support problem is avoided if the conditioning variables vary only gradually, since one must only apply matching sequentially for adjacent time periods in stead of once for periods that far off. We relax Assumption 5 in the following way:

Assumption 6 $\forall\ k, \forall\ x_k :\ Prob(X_{k-1} = x_k) \neq 0$

We therefore only require that, for any realisation x_k at time k, we can find with a strictly positive probability the same realisation in the preceding month $k-1$. If satisfied together with Assumptions 3 and 4, the following proposition shows how one can identify ATT_k.[29]

Proposition 5 [30] *If Assumptions 3, 4 and 6 hold then:* $\forall\ k \geq k^*$:

$$ATT_k =$$

$$\sum_{j=k^*+1}^{k} E\left[E(Y_j|D=1, X_j = x_j) - E(Y_{j-1}|D=1, X_{j-1} = x_j)|D=1\right] +$$

$$E\left[E(Y_{k^*}|D=1, X_{k^*} = x_{k^*}) - E(Y_{k^*-1}|D=0, X_{k^*-1} = x_{k^*})|D=1\right]$$

The proposition suggests estimating ATT_k by a sum of incremental impacts in the successive periods between the start of the treatment period and the moment k at which the eventual (cumulative) impact is evaluated. However, since the outcome variable is in generally positively serially correlated, this procedure will come at the cost of inflating the standard error of the estimated impact.[31] We therefore apply the procedure of incremental matching only if necessary, i.e. on the treated observations outside the region of common support if direct matching is applied. Moreover, we reduce the number of terms in the sum by grouping time periods up to the point that the problem of common support shows up.

4.1.3 The Estimated Impacts according to Before-After Estimators

We aim at estimating the ATT_k of benefit exhaustion on the employment rate according to definition 1 for $k \in [-3, 14]$.[32] In this section we report the estimation results of the different before-after estimators. As a benchmark, we present the results on the basis of a linear adjustment function $g(.)$. Subsequently, we report the findings of direct and incremental matching estimators.

[29] In the empirical application we need not relax Proposition 3.

[30] see Appendix C for a proof.

[31] In the empirical application we estimate this standard error by bootstrapping.

[32] In a sensitivity analysis we contrasted some of the results to the ones obtained by taking the employment rate according to definition 2 as the outcome of interest. These findings can be obtained from the authors upon request, but they do not modify any of our conclusions.

Table 3: Treatment Effects in the Pre-Treatment Period [a]

k [b]	Number of matched observations	Linear adjustment function		Direct Matching	
		ATT_k	SE^c	ATT_k	SE^c
-5	825	-0.008	0.006	-0.006	0.008
-6	825	-0.003	0.006	0.007	0.008
-7	823	-0.009	0.007	-0.003	0.009
-8	822	-0.020	0.008	-0.015	0.009
-9	821	-0.011	0.007	0.007	0.010
-10	820	-0.011	0.008	0.000	0.010
-11	820	-0.013	0.010	0.006	0.010
-12	819	-0.027	0.010	-0.015	0.011
-13	814	-0.009	0.009	0.007	0.009

[a] k =-4 is the reference
[b] Number of months of entitlement to UI left prior to the benefit exhaustion = -k
[c] Bootstrapped standard error (1000 replications).

Implementation and Results on the Basis of a Linear Adjustment Function

We estimate the linear adjustment function on the basis of a linear regression of the outcome variable at $k = -4$ on the above-defined unemployment rate and duration at that instant. In principle we could augment the sample by adding the prior periods, but we did not do so as the sample is incomplete before $k = -4$ (see Section 3).

Before turning to the estimates of the average treatment effects, we discuss the results of the testable implications of our assumptions as formalised in Proposition 1: the pre-treatment (adjusted) outcomes may not differ over time. We contrast the outcome at $k = -4$ to those in the interval $[-13, -5]$. Again note that the sample size ($N = 826$) gradually decreases from 825 to 814. We do not go beyond $k = -13$, since the sample size reduces more quickly after that point: this may invalidate the test. The findings are reported in the third column of Table 3. In all the periods tested we find negative values for ATT_k. Moreover, they are rather small. Only in two periods, the impact is statistically significant below the 5% level ($k = -8$ and $k = -12$). We will see below that this problem is induced by a misspecified adjustment function.

In the 4th column of Table 4 and Figure 4 we report the estimated ATT_k on the basis of the linear adjustment function for all $k \in [-3, 0]$ and for k equal to 3, 6, 12 and 14. We also contrast these results with those obtained by a "naive" before-after estimator. The latter estimator does not control for the time-varying variables and boils down to a matching estimator in which each woman in the "after period" is matched to herself in the "before period" (Heckman *et al.*, 1999). It is implemented by simply taking the difference between the average outcome at k and in the the first pre-treatment month ($k = -4$).

It turns out from Figure 4 that both estimators yield very similar results. This is because the two time-varying variables affect the outcome in opposite directions. First, during the observation period the economic conditions improve and this is

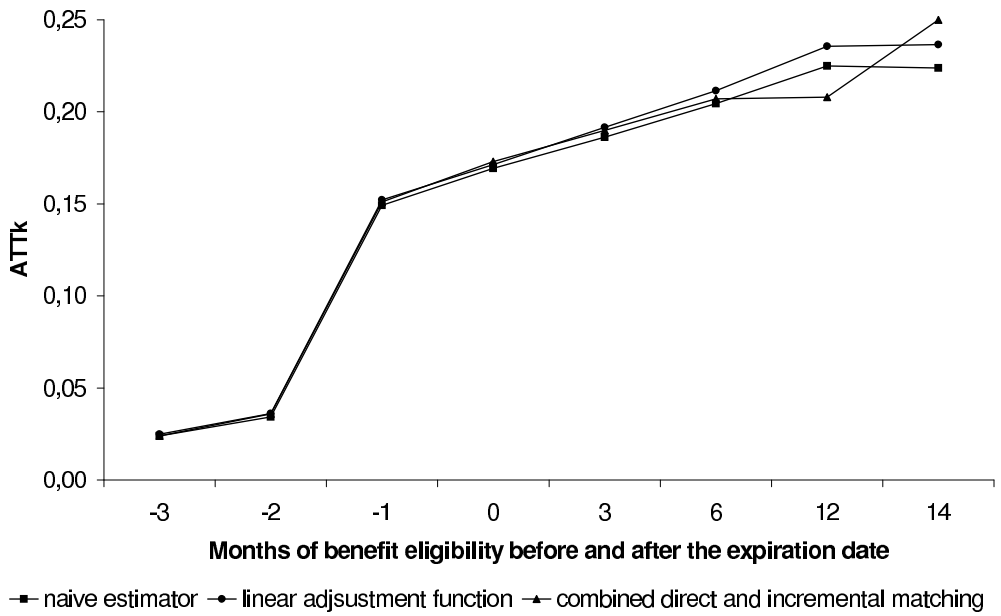
Table 4: Different Before-After Estimators of ATT_k

k	naive estimator		linear adjustment function		combined direct and incremental matching	
	ATT_k	SE [a]	ATT_k	SE [b]	ATT_k	SE [b]
-3	0.024	0.005	0.025	0.006	0.024	0.009
-2	0.034	0.006	0.036	0.008	0.036	0.010
-1	0.149	0.012	0.152	0.012	0.151	0.013
0	0.169	0.013	0.171	0.014	0.173	0.015
3	0.186	0.014	0.192	0.014	0.190	0.016
6	0.204	0.014	0.211	0.015	0.207	0.016
12	0.225	0.015	0.236	0.016	0.208	0.033
14	0.224	0.015	0.237	0.017	0.250	0.033

[a] Analytical standard error: $SE(ATT_k) = \sqrt{\frac{ATT_K * (1 - ATT_K)}{N}}$

[b] Bootstrapped standard error (1000 replications).

Figure 4: Different Before-After Estimators of the ATT_k



reflected in a decrease in the local unemployment rates. On the other hand, the cumulative unemployment duration can, by construction, never decrease. Apparently, these two effects cancel each other out.

Implementation and Results of the Direct and Incremental Matching Estimators

We first explain the benchmark propensity score matching method that we apply in this section. Subsequently, we discuss how the test for the absence of any impact in the pre-treatment period performs on the basis of a direct matching estimator. We then focus on a number of problems and choices to be made in the estimation of the ATT_k for $k = 14$. On the basis of this discussion we select a method of estimation that we implement to estimate the ATT_k for the same set of values of k as above.

The propensity scores are estimated using a procedure proposed by Dehejia and Wahba (2002) [33]. This procedure aims at finding the correct specification of the propensity score that satisfies the "balancing property" (Rosenbaum and Rubin, 1983). Once the propensity score is estimated, we must select a matching algorithm. Our benchmark choice is based on a method of "blocking" in which the matching occurs over a number of blocks into which the support of the estimated propensity score has been divided (see e.g. Rosenbaum and Rubin (1984) and Dehejia and Wahba (1999, 2002)). The same blocks, as those used for the specification test of the propensity score can be used. The standard error of ATT_k is calculated by a non-parametric bootstrap.

We choose the blocking estimator mainly for its simplicity and because empirical practice with matching estimators suggests that different matching algorithms generally tend to produce similar results. We nevertheless analyzed the sensitivity of the results to this particular choice of the matching algorithm.

As for the method based on the adjustment function, we first test for the temporal stability during the pre-treatment period of the outcome variable conditional on the time-varying explanatory variables (see Proposition 3). The results are reported in column 5 of Table 3. In contrast with the previous method, no impact effect is now significantly different from zero at a 5% level. This confirms that the adjustment function is misspecified.

We now turn to the estimation of the treatment effects. We first tried to estimate ATT_{14} by directly matching observations at $k = 14$ with those in the pre-treatment period at $k = -4$. On the basis of the above-mentioned benchmark procedure we could not find any balanced specification of the propensity score. [34] We suspect that this failure is the consequence of the true specification not being sufficiently smooth to be captured by a logit transformation of a polynomial in the continuously valued explanatory variables. If the propensity score varies abruptly in regions that lack common support, a propensity score that is a function of continuously valued variables cannot be balanced in these regions. We therefore devised a procedure in which we categorize the explanatory variables in a discrete

[33] see appendix E for a description.

[34] We tried specifications with several higher order and interaction terms as well as ones in which we added the time-constant explanatory variables.

number of intervals as to identify these regions. The procedure consists in five steps:

1. Split up the support of each of the variables in 5 equal-sized intervals;
2. To ensure balancing within each interval, test whether the mean of the variable within each interval does not significantly differ between the $k = -4$ group and the $k = 14$ group;
3. If the test in step 2 is rejected for some interval, split the interval in two and go back to the beginning of step 2;
4. After convergence of this procedure, construct cells on the basis of all possible combinations of the retained intervals of these variables;
5. Identify the region of common support by retaining only cells that contain observations for both $k = 14$ and $k = -4$.

On the basis of this procedure we identified 249 women, representing 30% of the sample, for whom the propensity score lacks common support. If we now apply the above-mentioned benchmark matching procedure on the 577 remaining women for whom we identified a common support, we no longer have any problems in finding a specification for the score that is balanced. The estimate of ATT_{14} of this group is reported in the first line of Table 5. The treatment effect is slightly higher than the one obtained by the methods of the previous section. However, this result may be biased, because it neglects a sizeable proportion of treated women. We therefore considered the procedure of incremental matching to re-incorporate the women outside the common support in the estimation procedure.

The second line of Table 5 reports the result of the incremental matching procedure in which we take the sum of incremental impacts in every month between $k = -4$ and $k = 14$. The advantage of this method is that the benchmark procedure works and that we face nowhere a common support problem. The cost is, as anticipated, a huge increase in the standard error, inflating from 0.024 to 0.282!

In order to increase the precision, we modify the method in two respects. First, we apply the incremental matching estimator only on the 249 women who were outside the common support. The total ATT_{14} is then found by taking a weighted average[35] of the impact of the latter group and the impact of the women within the common support, as estimated by the direct matching method. Second, rather than taking the sum of every monthly increment, we reduce the number of terms in the sum by increasing the length of each increment in the sum. To determine the length of each increment, we choose the maximal length for which the propensity score satisfies the balancing property without having to introduce higher order terms or interactions of the explanatory variables. In the empirical application this allows us to reduce the number of increments from 18 to 4. Note that the standard error of this procedure is estimated by bootstrapping.

The third line of Table 5 reports the point estimate of ATT_{14} for the previously neglected group. It clearly lower than the point estimate for the retained group:

[35] The weights are taken to be proportional to the size of each sub-sample.

0.198 versus 0.272. The weighted average estimate of ATT_{14} for the *total* sample, reported in the fourth line, is therefore slightly smaller (0.250). However, more importantly, the standard error of this new estimate is of a reasonable magnitude (SE = 0.033) as compared to the one obtained by the above reported incremental matching on the total sample (SE = 0.282).

Table 5: Estimates of ATT_{14} using Direct and Incremental Matching Methods

		treatment effect	standard error	N treated
1	direct comparison [a]	0.272	0.024	577
2	IATT (1) [b]	0.224	0.282	826
3	IATT (2) [c]	0.198	0.094	249
4	direct + IATT (2) [d]	0.250	0.033	826
5	kernel (r.o.t.) [e]	0.217	0.033	826
6	kernel (r.o.t.* 10) [f]	0.232	0.032	826

[a] direct comparison for the individuals on the common support
[b] sum of incremental impacts for the whole sample
[c] sum of incremental impacts for the individuals out of the common support
[d] weighted average of (1) and (2)
[e] same as (4), but using a kernel matching method, with bandwidth $h = 0.02$
[f] same as (4), but with $h = 0.2$

Before applying the proposed estimator of ATT_k for k's smaller than 14, we report a small sensitivity analysis with respect to the choice of the matching algorithm for ATT_k. We chose to contrast the blocking algorithm to a kernel matching estimator (see Heckman *et al.*, 1998). The basic difference between the kernel and the blocking method lies in the weighting function. Blocking gives a uniform weight to each control unit within a block (and zero if it is outside this block). The kernel method attributes a weight that is a decreasing function of the distance between the propensity score of treated and control units. The rate of decrease depends on the particular choice of the kernel function and (more importantly) on the smoothing parameter.

In this application we have chosen the Epanechnikov kernel function and tested how the results are affected by two different choices of the smoothing parameter. The first choice corresponds to the one obtained by Silverman's rule of thumb (Silverman, 1986). As this value was quite small, the second value was chosen to be 10 times higher. The estimation results according to these methods are reported in the two last lines of Table 5. Both methods yield lower estimates of the ATT_{14} than did the blocking method. The differences are, however, not significant: the kernel estimates are inside the 95% confidence interval of the blocking estimates. Note that the estimates obtained with a larger bandwidth come closer to the blocking estimates. This is because the (relative) weights are distributed more uniformly by using a larger bandwidth, thus coming closer to the weighting scheme of the blocking method, where all the observations of the control group receive equal weight.

We finally report in Table 4 the results of the combined direct and incremental matching estimator for other treatment months k. The benchmark blocking algorithm was again applied. We observe that in this empirical application the

results hardly depend on the method by which one accounts for time-varying determinants of the outcome variable.

On the basis of these findings, we conclude that benefit exhaustion does significantly enhance the probability of employment: by 17.3 percentage points in the month in which benefits are withdrawn, gradually growing to 25 percentage points 14 months afterwards. Moreover, already in the month in which one is notified the employment probability is 2.4 percentage points higher. Subsequently, the impact jumps up from 3.6 percentage points two entitlement months before the expiration to 15.1 percentage points one month before. This demonstrates that anticipation effects are important.

4.2 A Sensitivity Analysis: Standard Propensity Score Matching and IV

One may be suspicious of the validity of the before-after estimator since it requires the strong identifying assumption that all systematic time-varying effects can be captured by a number of observed time-varying variables. In order enhance the credibility of the estimation result, we therefore test whether similar results could be attained on the basis of different identifying assumptions. We will consider two alternative methods: standard propensity score matching on observed explanatory variables and an Instrumental Variable (IV) method.

In contrast with the before-after estimator, the alternative methods are not only based on the sampled treatment group, but also on the control group, i.e. the group of women for whom benefits did not (yet) exhaust at the survey date. However, the sampling scheme complicates an analysis that uses both groups, since the sampling criteria differ between these groups (see Section 3). Women in the control group are selected among those who received unemployment benefits during at least one day in March 1997. In contrast, a sizeable proportion (55%) of the treated women were not unemployed in that month. There is clearly a lack of common support.

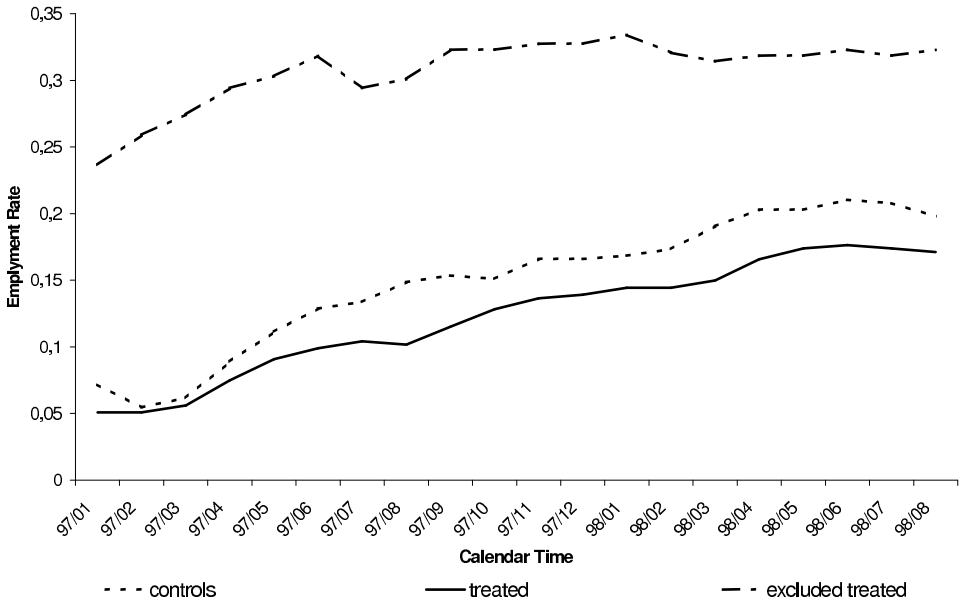
In order to resolve the problem, we exclude the 462 treated women who were not unemployed in March 1997 from the analysis. This eliminates women who are attached to the labour market. The excluded and retained treated do not differ very much with respect to their observed individual characteristics.[36] However, they differ quite a lot as far as their labour market history is concerned. The excluded treated have on average a slightly lower cumulated unemployment duration than the retained ones (respectively 52 and 56 months) but, in particular, their last unemployment spell is shorter (respectively 19 and 36 months). The excluded treated therefore seem to have a more volatile employment history: between January 1992 to March 1997 they experienced 3.3 unemployment spells on average, whereas the retained group only 2.8. The lack of recent employment experience of the retained women as compared to the disregarded group also shows up in Figure 5. In this figure the average employment rates of the different groups are plotted between January 1997 and August 1998[37]. The employment rate of the excluded

[36] A description of the observed individual characteristics and the labour market histories of the the various sub-samples is reported in Table 8 of Appendix D.

[37] The employment status corresponds to the one reported in the survey.

treatment group is much higher than the one of the retained group. The latter is now even everywhere consistently lower than the employment rate of the control group.

Figure 5: Monthly Employment Rates of Treated and Controls



A first consequence of this sample restriction is that the effects of benefit exhaustion prior to March 1997 are difficult to interpret, since they are conditional on a future event: being unemployed in March 1997. As in the treatment sample benefits expire between March and June 1997, we decided to disregard the anticipation effects in the estimation. In this section the ATT_k are therefore only estimated for $k \geq 0$.

A second consequence of this restriction is that we can no longer compare the results obtained by the alternative methods with the ones of the before-after estimator reported in the previous section. The comparisons reported below will therefore refer to the before-after estimator as applied on the restricted treatment sample.[38]

4.2.1 The Standard Matching Method

We first consider the standard matching method. The method consists in matching members of the retained treatment and control sample on the basis of the retained control variables reported in Table 2 of Section 3. To avoid an endogeneity bias, we condition on the realisations of the time varying variables at the beginning of

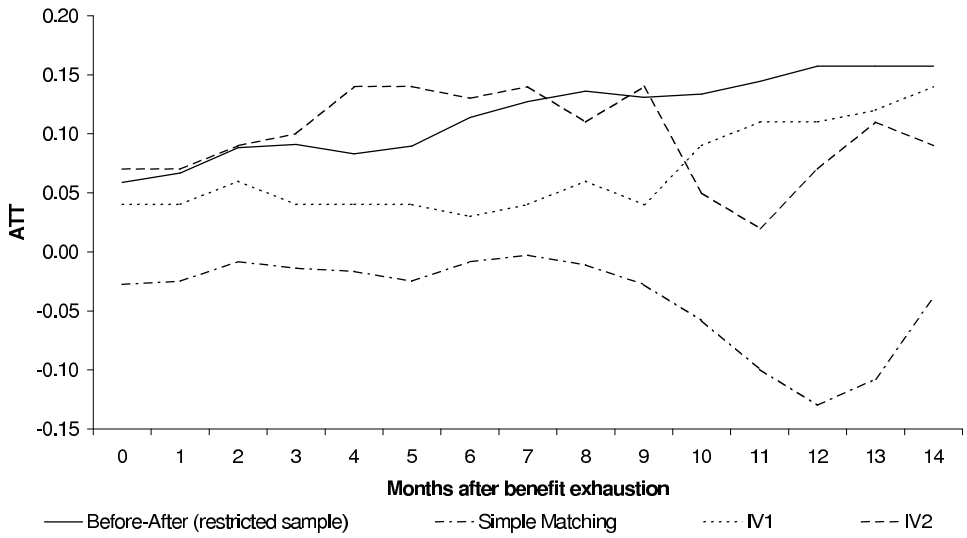
[38] Note that on this restricted sample, we did not have any problem of common support. Consequently, we did not have to rely on the incremental matching method, but could use direct matching instead.

March 1997, i.e. prior to the first benefit exhaustion that we observe.[39]

We used the benchmark procedure described in the previous section to find a correctly specified propensity score. However, in contrast to the before-after estimator, we applied the nearest-neighbour matching algorithm with replacement in stead of the blocking estimator. The reason is the following complication. Since we only condition on a set of time-constant explanatory variables, we must control for time-varying effects by contrasting the outcomes of treatments and controls at the same calendar month. However, since treatments do not all occur on the same calendar time – benefit exhaust between March and June 1997 – this means that we should evaluate the outcome of a control at a different calendar time according to the treatment unit to whom she is matched. This differentiation is feasible if the treatment units are matched to only one control, as is the case with nearest-neighbour matching. For matching algorithms, such as blocking or kernel, in which treated units are matched with several controls, this is more difficult. However, as results are in general not very sensitive to the choice of the matching algorithm, we do not believe that this choice is restrictive.

The results of the impact are reported in table 6 and contrasted to the before-after estimator as applied on the restricted sample in Figure 6. Over the whole period, the impact is negative in a range between -11% and 0%. However, the standard errors are large, so that we cannot reject the hypothesis that benefit exhaustion had no impact on employment probabilities. On the other hand, the before-after estimate of the impact is not contained in the 95% confidence interval. In the next section we will verify whether this is due to some selection on unobservables not accounted for in the standard matching estimator.

Figure 6: Alternative Matching Estimators and the Before–After Estimator



[39] De Lathouwer *et al.* (2003) condition on the income of the partner at the time of survey, after benefits have exhausted. If the partner reacts to the income loss, e.g. by increasing the number of working hours, then the employment status of the treated women is negatively related to this income. This induces a downward bias in the estimated impact.

Table 6: Various matching estimators for ATT_k

k [a]	Before-After [b]	Simple Matching		IV1		IV2	
	ATT	*ATT*	*SE* [e]	*ATT*	*SE* [e]	*ATT*	*SE* [e]
0	0.06	-0.03	0.05	0.04	0.07	0.07	0.24
1	0.07	-0.02	0.05	0.04	0.07	0.07	0.26
2	0.09	-0.01	0.06	0.06	0.08	0.09	0.28
3	0.09	-0.01	0.06	0.04	0.08	0.10	0.29
4	0.08	-0.02	0.06	0.04	0.08	0.14	0.29
5	0.09	-0.02	0.06	0.04	0.08	0.14	0.29
6	0.11	-0.01	0.06	0.03	0.09	0.13	0.30
7	0.13	0.00	0.07	0.04	0.09	0.14	0.32
8	0.14	-0.01	0.07	0.06	0.09	0.11	0.32
9	0.13	-0.03	0.07	0.04	0.09	0.14	0.32
10	0.13	-0.06	0.07	0.09	0.09	0.05	0.33
11	0.14	-0.10	0.07	0.11	0.09	0.02	0.34
12	0.16	-0.13	0.07	0.11	0.09	0.07	0.34
13	0.16	-0.11	0.07	0.12	0.09	0.11	0.35
14	0.16	-0.04	0.07	0.14	0.09	0.09	0.35

[a] Months after benefit exhaustion
[b] Direct matching applied to the restricted sample
[c] IV–matching using a "simple" specification (c.f. discussion in Section 4.2.2)
[d] IV–matching using a "complete" specification (c.f. discussion in Section 4.2.2)
[e] Analytical standard error (c.f. Appendix F)

4.2.2 IV as a matching estimator

According to legislation UB can only be withdrawn if the cumulative unemployment duration exceeds a certain threshold. This threshold depends on the sex, the age class and the place of residence. These rules assign recipients of UB into an eligible and non-eligible population. Benefits will not expire for all eligible women, since there exist several grounds for lodging an appeal (see Section 2). Nevertheless, as a consequence of these rules, the probability of UB exhaustion jumps up, once assigned to the eligible population. We exploit this variation in the probability of treatment to identify the effect of UB expiration.

The variation in the probability of treatment induced by the legislation is not completely exogenous, however: the above-mentioned determinants of assignment to the eligible population are not exogenous to the outcome variable, the probability of employment. We therefore first need to purge this variation of its component that is correlated with the employment rate. This is realised on the basis of a "Regression Discontinuity Design" (see e.g. Hahn *et al.* (2001)). We assume that the relationship between the employment probability and the determinants of eligibility is smooth. In contrast, the relationship between these determinants and the eligibility status is discrete: the duration must exceed a threshold and this threshold in turn varies discontinuously as the age crosses the boundary of an interval, or as the place of residence changes. Consequently, there will remain some variation in the eligibility status once we control for its determinants by means of

continuous variables. Moreover, this variation is exogenous in that it is no longer correlated with the outcome variable. This defines an instrumental variable.

We chose the following continuous control variables: the cumulated unemployment duration since April 1992, as calculated from the administrative files; the age in years at the sampling date[40]; the female local unemployment rate as reported in the monthly statistics of the National Unemployment Office (RVA/ONEM). The unemployment rate was taken to eliminate the correlation between the employment rate and the place of residence. The unemployment duration and rate were both evaluated at the end of February 1997.

The eligibility status depends on the cumulative unemployment duration since the first moment at which the worker became eligible to UI. As already mentioned in Section 4.1, the unemployment duration, as calculated from the available administrative files, underestimates this variable. We therefore determine the eligibility status on the basis of the estimated adjusted duration, as explained above. The eligibility status is determined in June 1997. By this choice, we ensure that all women for whom we observe the benefit exhaustion belong indeed to the eligible population: all observed expirations occur before the end of June 1997. This choice avoids determining a different eligible population according to the date at which benefits exhaust. On the basis of this definition we identify 632 eligible women and 118 ineligible.

The discrete nature of the outcome variable complicates the implementation of an IV estimator. For, the Two Stage Least Squares estimator is then inconsistent (see e.g. Wooldridge (2002, p.478)). To avoid this problem, we follow the suggestion of Heckman *et al.* (1999) and Ichimura and Taber (2001) to implement IV as a matching estimator. This choice has the further advantage of being non-parametric. It avoids arbitrary parametric assumptions on both the distribution of the binary outcome and on the functional relation of the outcome with the explanatory variables.

The eligibility status, Z is a discrete IV, conditional on the above mentioned continuous variables, X, fixed at their realisations x at the end of February 1997. Y_{pkt} denotes the potential outcome ($p = 0, 1$) k months after benefit exhaustion at calendar time t.[41] Following Heckman *et al.* (1999) we may write the potential and observed outcome respectively as follows:

$$Y_{pkt} = g_{pkt}(X) + U_{pkt} \tag{4}$$

$$Y_{kt} = g_{0kt}(X) + D[\Delta_{kt}(X) + U_{1kt} - U_{0kt}] + U_{0kt} \tag{5}$$

where $\Delta_{kt}(X) \equiv g_{1kt} - g_{0kt}$. In this notation the ATT_k is defined by the following expression:

$$ATT_k \equiv E\left[\Delta_{kt}(X) + E(U_{1kt} - U_{0kt}|D = 1, X)|D = 1\right] \tag{6}$$

[40] In March 1997 for the controls and between March and June 1997 for the treated women.

[41] Since UB do not all exhaust at the same calendar time, but between March and June 1997, we distinguish k and t.

where the outer expectation is over the distribution of $X|D = 1$ and, for each k, over the four different t's , corresponding to the four months - March to June 1997 - in which benefits are observed to exhaust.

The next three assumptions allow us to identify ATT_k by IV as a matching estimator.

Assumption 7 $E(U_{0kt}|X, Z) = E(U_{0kt}|X)$

This is the standard IV condition that the the eligibility status may not be correlated with the no-treatment residual after conditioning on the above-mentioned continuous variables, X.

Assumption 8 $E(U_{1kt} - U_{0kt}|X, Z, D = 1) = E(U_{1kt} - U_{0kt}|X, D = 1)$

This assumption implies that the average idiosyncratic impact of benefit exhaustion on the treated may not depend on the IV. In our empirical application this assumption is not restrictive, since the IV corresponds to the eligibility status and therefore women are by definition only treated if $Z = 1$. In other words, there are "compliers" and "never-takers", but no "always-takers". As a consequence, the Local Average Treatment Effect (LATE) identified by the IV estimator (Imbens and Angrist, 1994) coincides with the ATT in this empirical application.

The third assumption requires that for any X the probability of treatment differs according to the realisation of the IV. Since here the IV corresponds to the eligibility status, this assumption is trivially satisfied:

Assumption 9 $P_1(X) \neq P_0(X) = 0$
where $P_z(X) \equiv Pr(D = 1|X, Z = z)$ *for* $z = 0, 1$.

If the above assumptions hold, then the following proposition suggests a feasible IV estimator of the ATT_k:

Proposition 6 [42] *If Assumptions 7-9 hold then:*

$$ATT_k = E\left[E\left[\frac{Y_{kt}}{P_1(X)}|X, Z = 1\right] - E\left[\frac{Y_{kt}}{P_1(X)}|X, Z = 0\right]|Z = 1\right]$$

where the outer expectation is over the distribution of $X|Z = 1$ *and over the* t*'s.*[43].

This is a Wald–type estimator. By the above arguments, conditional on X, the systematic difference in the outcome between the eligible ($Z = 1$) and the ineligible ($Z = 0$) must be induced by the benefit exhaustion. Since benefits do not exhaust for every eligible woman (c.f. the conditions of appeal in Section 2), we have to weigh this impact by the fraction, $P_1(X)$, of women among the eligible population for whom benefits effectively expire.

The above estimator is not operational, since it conditions on a vector of continuous variables, X: one cannot find women with both values of Z with exactly

[42] see Appendix C for a proof.

[43] If $D = 0$ and $Z = 1$, there is no natural starting point for calendar time. We arbitrary set t equal to March 1997 if $k = 0$.

the same realisations of X. To resolve this problem, Ichimura and Taber (2001) generalise the result from Rosenbaum and Rubin (1983) obtained for standard matching methods to IV matching. These authors show that, rather than conditioning directly on X, one may condition on the propensity of eligibility, $Q(X)$, in stead. Using this result they propose to match on the propensity score of eligibility in stead of on the propensity score of treatment. If we apply the nearest-neighbor matching routine with replacement, this results in the following operational estimator of ATT_k:

$$ATT_k = \frac{1}{N_{Z_1}} \sum_{i \in Z_1} \left[\frac{Y_{kti}}{\hat{P}_1(x_i)} - \frac{Y_{ktj(i)}}{\hat{P}_1(x_{j(i)})} \right] \tag{7}$$

where Z_z $(z = 0, 1)$ denotes the set of respectively ineligible $(z = 0)$ and eligible $(z = 1)$ women, N_{Z_1} the number of eligible women, $\hat{P}_1(x)$ is the estimated propensity score of treatment conditional on $Z = 1$, $j(i) = argmin_{j \in Z_0}[\hat{Q}(x_i) - \hat{Q}(x_j)]^2$, where $\hat{Q}(x)$ is the estimated propensity score of $Z = 1$ conditional on the realised explanatory variables x. Note that for $Z = D$ this boils down to a standard nearest-neighbor matching estimator.

We follow the benchmark procedure of Dehejia and Wahba (2002) described above and in Appendix E to obtain a balanced specification of the estimated propensity scores. Note, as for the standard matching method, we adjust the proposed procedure slightly as to account for the calendar time variation t in the outcome variable. The standard errors are calculated analytically, assuming that the propensity scores are fixed. [44]

We implement two variants of this IV estimator. "IV1" controls only for three explanatory variables: the unemployment rate and duration, and the age. This set of conditioning variables is in principle sufficient, since, by the above arguments, all remaining variation should be random. However, if our sample were an unlucky draw, this assumption could be violated. "IV2" We therefore tested whether the findings uphold if we include in addition in the specification of the scores all the remaining explanatory variables, i.e. those retained for the standard matching estimator.

The estimation results are reported in the last columns of Table 6 and in Figure 6. For both variants they are much closer to the before-after than to the standard matching estimates. The IV2 estimates are mostly closer to the before-after than IV1, suggesting that some bias remains present in the IV1 estimator. However, the IV2 estimator seems to suffer from problems of over-specification: standard errors more than triple in magnitude! Nevertheless, despite their imprecision, the IV estimates corroborate the findings of the before-after estimator.

5 Conclusion

This paper aimed at estimating the causal impact of UB exhaustion on the probability of employment of a specific category of long-term unemployed women in Belgium. To attain this objective we had to surmount a number of methodological

[44] see the Appendix F for a derivation of the formula.

problems. Before summing up the empirical findings, we briefly summarize the proposed resolution to the main problems.

First, we have shown that the fulfillment of the identifying assumptions of a before-after estimator of a causal impact may crucially depend on the definition of the time-scale, i.e by using the number of months of benefit receipt prior to the UI exhaustion date in stead of calendar time. Second, for the implementation of the before-after estimator, we proposed a non-parametric matching method to purge the outcome variable of time-varying determinants. As to solve a problem of common support that was bound to occur, we developed a method that did not require dropping any observations: the method of "incremental matching". This method boils down to taking the sum of incremental impacts over a number of sub-periods over the period of interest. Finally, as to overcome the inconsistency of the IV estimator in the presence of a discrete outcome variable, we implemented IV as a matching estimator, as suggested by Heckman *et al.* (1999) and Ichimura and Taber (2001).

The reader may question why we did not analyze the effect of benefit exhaustion by hazard regression models, as in the literature. A first answer is that matching estimators are non-parametric and therefore less sensitive to specification errors[45]. However, the main reason is the complexity of sampling scheme that, as mentioned above, combines a stock sample for the controls and a flow sample into benefit exhaustion for the treated women. A drawback of our choice is that it makes it more difficult to compare the magnitude of the estimated impact to previous findings and to test the predictions of theory which are also phrased in terms of hazard rates from unemployment to employment and not in terms of employment probabilities. For instance, job search theory predicts that the hazard rate should jump up at the moment of notification, should then gradually increase up to the expiration date and remain constant afterwards. Note also that the hazard from employment to non-employment should follow the reverse pattern (up to the point at which the worker is entitled to UB again). Since the employment probability is the sum (integral) of past hazard rates to employment times the probability of surviving in employment up to the considered moment, this pattern is smoothed: we should therefore observe the employment rate to increase more gradually, to continue this upward movement beyond the moment at which benefits expire and, subsequently, to converge gradually to a higher level. This pattern corresponds exactly to our findings.

These findings can be summarized as follows. On the basis of the before-after estimator we conclude that in Belgium the exhaustion of unemployment benefits of long-term unemployed female workers has an important significant positive impact on the probability of employment. From the moment at which the worker is notified, three months before expiration, the probability of employment rises. One month before benefits expire it already attains a level that is 16 percentage points higher than in the absence of benefit exhaustion and afterwards the impact rises more gradually up to 25 percentage points 14 months after the end of the entitlement period. If benefits had not expired, the employment probabilities at

[45] In most cases a mixed proportional hazard (MPH) assumption is required. This can only be relaxed if one has data on multiple spells for the same individual (Abbring and van den Berg, 2003).

these moments would have been respectively of 2.5% and virtually zero (according to our estimations).

For the interpretation of the results it is important to realise that the treatment sample was restricted to women whose benefits effectively expired. The impact is therefore underestimated, since it does not take into account that women could escape the benefit exhaustion by leaving unemployment for a full time job lasting more than two years. After such an employment period workers regain entitlement to unemployment benefits and the unemployment duration clock is reset to zero. On the other hand, the restriction of the analysis to those women whose benefits were effectively withdrawn necessarily implies that all employment found prior to the moment of benefit exhaustion must have lasted less than two years and can therefore be qualified as "temporary".

There is a second reason why the before-after estimator underestimates the impact on the employment probability. Since we condition on the elapsed unemployment duration in the estimations, we disregard the indirect positive effect of benefit exhaustion on the employment rate via its negative impact on the rate of increase of the unemployment duration. However, there is no reason to believe that this indirect effect is large.

The before-after estimator has been compared to both a "standard" and an IV matching estimator, be it on a smaller sample, as a consequence of a lack of common support between the treatment and the control sample. Even if the IV matching estimator was imprecise, it yielded very close estimates of the ATT to the before-after estimator on the corresponding sample. This corroborates our findings. On the other hand, the standard matching estimator that only corrects for selection on observables consistently yielded negative impact estimates and therefore performed very badly in this empirical application.

We conclude that our findings are in the line of the more recent studies, mentioned in the introduction that find important disincentive effects of UB. However, these findings concern a very specific sub-population of unemployed women in Belgium. We do not know whether these results generalise to other unemployed workers. Moreover, we did not evaluate the quality of the realised employment (as measured by the wage or other employment characteristics such as the fraction of full-time employment or of permanent contracts, etc.). This requires further research.

Appendix

A Constructing an Instrument

The IV is a binary variable indicating whether an individual has crossed the threshold duration, as defined by the legislation, or not. To make things more precise, let UD_i be the unemployment duration and let TD_i be the threshold duration for individual i. This allows us to define the instrument as:

$$\begin{aligned} Z_i &= 1 \quad if \quad \delta_i > 0 \\ Z_i &= 0 \quad else \end{aligned}$$

where $\delta_i \equiv UD_i - TD_i$ is the distance between the effective unemployment duration and the threshold duration. Note that a major implication of this definition is that we must have $Z_i = 1$ for any treated individual. It is also crucial to realize that the reverse is not true: for a treated individual, Z_i can take both zero or one. Indeed, crossing the threshold duration is a necessary condition for benefit exhaustion, but not a sufficient one. Nevertheless, in our data set a problem arises. The individual unemployment duration is only partially observed. More precisely, it is left-censored in January 1992. Letting UD_{0i} and δ_{0i} the observed values of δ_i and UD_i, we have

$$\delta_{0i} = UD_{0i} - TD_i$$

for some (or even all) individuals. As an immediate consequence, Z_i is no longer equal to one for all the treated. We therefore propose the following correction procedure. First, we select the treated for which δ_{0i} takes the correct sign (i.e. is negative). In a second step, a regression model is specified for those individuals:

$$\log(-\delta_{0i}) = X_i'\beta + \epsilon_i$$

where X_i is a set of individual characteristics, including a constant term and β a set of coefficients to be estimated (by OLS). Next, we use the the estimated values of β to obtain the predicted value of δ_{0i} for the whole sample:

$$\hat{\delta}_{0i} = exp(X_i\hat{\beta} + \frac{1}{2}\hat{\sigma}_\epsilon^2)$$

where $\hat{\sigma}_\epsilon^2$ is the variance of the error term. With these elements in hand, the instrument is defined as follows:

$$\begin{aligned} Z_i &= 1 \quad if \quad UD_{0i} + \hat{\delta}_{0i} > 0 \\ Z_i &= 0 \quad else \end{aligned}$$

B Employment Status before January 1997

In this appendix we describe how the outcome variable prior to January 1997 was constructed. For this period, we observe monthly the number of days of unemployment benefit receipt. On the basis of this information, we propose an estimate of the employment status such as defined in the survey. To this purpose we estimate the relationship between the administrative and survey information during the period in which both sources are available: from January 1997 up to the sampling date (March–June 1997). Basically, we did the following. We have chosen a threshold number of days an individual spends in unemployment such that in January 1997, the fraction of individuals with a number of days in unemployment below this threshold is equal to the fraction of individuals which report themselves as being in employment. We chose January rather than some later month, because the number of days unemployed is a better predictor of the employment state the more remote one is to the benefit exhaustion date: the closer one is to this date, the less likely the non-receipt of UI corresponds with employment, since the worker is more likely to become inactive or to search employment without entitlement to benefits. As a consequence, the predicted employment status overstates the one reported in the survey, the closer one is to the expiration date (see Table 7 below).

1. Compute the fraction of individuals declaring themselves as being employed in January 1997. Denote this fraction by f_E.

2. Compute for each individual the fraction of time he has spent in unemployment (or sickness) in January 1997 (according to the administrative files). Call this fraction f_U.

3. Use a logit model to predict the reported employment status in January 1997 as a function of f_U, and a vector of individual attributes. Let us denote this prediction $\hat{E}$.

4. Obtain the empirical survivor function of $\hat{E}$, $S(\hat{E})$.

5. Find the value of $\hat{E}$, such that $S(\hat{E}) = f_E$. Call this fraction f^*.

6. Define the employment status according to the administrative files E_A as follows:
   $E_A = 1$ if $f > f^*$ and $E_A = 0$ else

We computed a different threshold fraction according to the sampling date (March-June 1997). In table 7 we compare employment status as computed from the administrative files (C) to the employment status as reported in the survey (R). The columns represent the different sampling dates, and the rows represent the months for which the two data sources overlap. By construction, the reported and the computed status are equivalent in January 1997. In the following months the predicted employment status is, as mentioned, always higher.

Table 7: Reported vs. Computed Employment Status

		March	April	Mai	June
January	R	0.18	0.15	0.14	0.12
	C	0.18	0.15	0.14	0.12
February	R	0.21	0.13	0.15	0.13
	C	0.24	0.16	0.16	0.12
March	R	0.22	0.13	0.17	0.13
	C	0.25	0.18	0.20	0.15
April	R		0.15	0.19	0.15
	C		0.21	0.25	0.21
Mai	R			0.21	0.15
	C			0.30	0.21
June	R				0.20
	C				0.24

Columns: Sampling date
Lines: Period of overlap
R: Reported employment status
C: Predicted employment status

C Proofs of the Propositions

Proposition 1 *If Assumptions 1 and 2 hold, then:*

$$\forall k' < k^* : E[A_{k'} - A_{k^*-1}|D = 0] = 0$$

Proof

$$\begin{aligned} E[A_{k'} - A_{k^*}|D = 0] &= E[A_{0k'} - A_{0k^*-1}|D = 0] \\ &= E[A_{0k'} - A_{0k^*-1}|D = 1] \quad \text{by Assumption 1} \\ &= 0 \quad \text{by Assumption 2} \end{aligned}$$

■

Proposition 2 $\forall k \geq k^*, \exists k' < k^* : ATT_k = E[A_k|D = 1] - E[A_{k'}|D = 0]$

Proof

$$\begin{aligned} ATT_k &\equiv E[Y_{1k}|D = 1] - E[Y_{0k}|D = 1] \\ &= E[A_{1k}|D = 1] - E[A_{0k}|D = 1] \\ &= E[A_{1k}|D = 1] - E[A_{0k'}|D = 1] + E[A_{0k'}|D = 1] - E[A_{0k}|D = 1] \end{aligned}$$

By Assumption 1 we have $E[A_{0k'}|D = 1] = E[A_{0k'}|D = 0]$ and by Assumption 2 we have $E[A_{0k'}|D = 1] - E[A_{0k}|D = 1] = 0$, which, using the definition of the observed outcome Y_k, yields the desired expression. ■

Proposition 3 *If Assumptions 3, 4 and 5 hold, then:*

$$\forall\ k' < k^*, \forall\ x_{k^*-1} : E[E[Y_{k'}|D = 0, X_{k'} = x_{k^*-1}] - E[Y_{k^*-1}|D = 0, X_{k^*-1} = x_{k^*-1}]|D = 0] = 0$$

Proof By 5 $E[Y_{k'}|D = 0, X_{k'} = x_{k^*-1}]$ exists and by using the definition of the observed outcome Y'_k one obtains:

$$\begin{aligned} &E[E[Y_{k'}|D = 0, X_{k'} = x_{k^*-1}] - E[Y_{k^*-1}|D = 0, X_{k^*-1} = x_{k^*-1}]|D = 0] \\ &= E[E[Y_{0k'}|D = 0, X_{k'} = x_{k^*-1}] - E[Y_{0k^*-1}|D = 0, X_{k^*-1} = x_{k^*-1}]|D = 0] \\ &= E[E[Y_{0k'}|D = 1, X_{k'} = x_{k^*-1}] - E[Y_{0k'}|D = 1, X_{k^*-1} = x_{k^*-1}]|D = 0] \quad \text{by Assumption 3} \\ &= 0 \quad \text{by Assumption 4} \end{aligned}$$

■

Proposition 4 *Assumptions 3, 4 and 5 imply that:*

$$\forall\ k \geq k^*,\ \exists\ k' < k^* :\ ATT_k = E[E[Y_k|D = 1, X_k = x_k] - E[Y_{k'}|D = 0, X_{k'} = x_k]|D = 1]$$

where the outer expectation is over the distribution of $X_k|D = 1$

Proof

$$ATT_k \equiv E[Y_{1k} - Y_{0k}|D = 1]$$

By the law of iterated expectations and adding and subtracting terms:

$$\begin{aligned} ATT_k = E[E[Y_{1k}|D = 1, X_k = x_k]|D = 1] - E[E[Y_{0k'}|D = 1, X_{k'} = x_k]|D = 1] \\ -E[E[Y_{0k}|D = 1, X_k = x_k] - E[Y_{0k'}|D = 1, X_{k'} = x_k]|D = 1] \end{aligned}$$

By Assumption 5, $E[Y_{0k'}|D = 0, X_{k'} = x_k]$ exists for each x_k. Moreover, by Assumption 3, we have $E[Y_{0k'}|D = 1, X_k = x_k] = E[Y_{0k'}|D = 0, X_{k'} = x_k]$ for all x_k. Finally, by Assumption 4 we have, $E[E[Y_{0k}|D = 1, X_k = x_k] - E[Y_{0k'}|D = 1, X_{k'} = x_k]|D = 1] = 0$. Using all this, we obtain:

$$ATT_k = E[E[Y_{1k}|D = 1, X_k = x_k] - E[Y_{0k'}|D = 0, X_{k'} = x_k]|D = 1]$$

Since $Y_k = DY_{1k} + (1 - D)Y_{0k}$, we obtain:

$$ATT_k = E[E[Y_k|D = 1, X_k = x_k] - E[Y_{k'}|D = 1, X_{k'} = x_k]|D = 1]$$

■

Proposition 5 *Assumptions 3, 4 and 6 imply that:* $\forall\ k \geq k^*$:

$$ATT_k =$$

$$\sum_{j=k^*+1}^{k} E\left[E(Y_j|D = 1, X_j = x_j) - E(Y_{j-1}|D = 1, X_{j-1} = x_j)|D = 1\right] +$$

$$E\left[E(Y_{k^*}|D = 1, X_{k^*} = x_{k^*}) - E[Y_{k^*-1}|D = 0, X_{k^*-1} = x_{k^*}]|D = 1\right]$$

Proof

$$ATT_k \equiv E[Y_{1k} - Y_{0k}|D = 1]$$

By adding and subtracting the same terms:

$$\begin{aligned} ATT_k = \\ = \textstyle\sum_{j=k^*+1}^{k} E[Y_{1j}|D = 1] - E[Y_{1j-1}|D = 1] + E[Y_{1k^*}|D = 1] - E[Y_{0k^*-1}|D = 1] \\ + \textstyle\sum_{j=k'}^{k^*-1} E[Y_{0j}|D = 1] - E[Y_{0j-1}|D = 1] + E[Y_{0k'}|D = 1] - E[Y_{0k}|D = 1] \end{aligned}$$

Using the law of iterated expectations, and the Assumption 6 to ensure the existence of the conditional expectations:

$$\begin{aligned} ATT_k = \\ \textstyle\sum_{j=k^*+1}^{k} E[E[Y_{1j}|D = 1, X_j = x_j] - E[Y_{1j-1}|D = 1, X_{j-1} = x_j]|D = 1] & \quad (1) \\ +E[E[Y_{1k^*}|D = 1, X_{k^*} = x_{k^*}] - E[Y_{0k^*-1}|D = 1, X_{k^*-1} = x_{k^*}]|D = 1] & \quad (2) \\ + \textstyle\sum_{j=k'}^{k^*-1} E[E[Y_{0j}|D = 1, X_j = x_j] - E[Y_{0j-1}|D = 1, X_{j-1} = x_j]|D = 1] & \quad (3) \\ +E[E[Y_{0k'}|D = 1, X_{k'} = x_k] - E[Y_{0k}|D = 1, X_k = x_k]|D = 1] & \quad (4) \end{aligned}$$

Using the fact that $Y_j = DY_{1j} + (1-D)Y_{0j}$, the terms in (1) can be rewritten as

$$\sum_{j=k^*+1}^{k} E[E[Y_j|D=1, X_j = x_j] - E[Y_{j-1}|D=1, X_{j-1} = x_j]|D=1]$$

Using Assumption 3 together withe the definition of Y_j, we can rewrite the terms in (2) as

$$E[E[Y_{k^*}|D=1, X_{k^*} = x_{k^*}] - E[Y_{k^*-1}|D=0, X_{k^*-1} = x_{k^*}]|D=1]$$

Finally, by Assumption 5, the terms in (3) and (4) are equal to zero. ■

Proposition 6 *If Assumptions 7-9 hold then:*

$$ATT_k = E\left[E\left[\frac{Y_{kt}}{P_1(X)}|X, Z=1\right] - E\left[\frac{Y_{kt}}{P_1(X)}|X, Z=0\right]|Z=1\right]$$

where the outer expectation is over the distribution of $X|Z=1$ and over the t's.[46].

Proof Using equation 5, Assumption 9 and the Law of Iterated Expectations, we can write:

$$E\left[\left[\frac{Y_{kt}}{P_1(X)}|Z=z, X\right]|Z=1\right] = E\left[\frac{g_{0kt}(X) + E(U_{0kt}|X, Z=z)}{P_1(X)}|Z=1\right] +$$

$$E\left[P_z(X)\frac{\Delta(X) + E(U_{1kt} - U_{0kt}|X, Z=z, D=1)}{P_1(X)}|Z=1, D=1\right]$$

Using Assumptions 7 and 8, and since $D=1$ only if $Z=1$, i.e. since $P_0 = 0$, this simplifies to

$$E\left[\left[\frac{Y_{kt}}{P_1(X)}|Z=z, X\right]|Z=1\right] = E\left[\frac{g_{0kt}(X) + E(U_{0kt}|X)}{P_1(X)}|Z=1\right] +$$

$$E\left[P_z(X)\frac{\Delta(X) + E(U_{1kt} - U_{0kt}|X, D=1)}{P_1(X)}|D=1\right]$$

Taking the difference of this expectation for $z=1$ and $z=0$ and Assumption 9, i.e. $P_0(X) = 0$, proves the Proposition. ■

[46] If $D=0$ and $Z=1$, there is no natural starting point for calendar time. We arbitrary set t equal to March 1997 if $k=0$.

D Sample Characteristics for the IV analysis

Table 8: Sample characteristics of the different groups

	C (N = 404)	T (N = 362)	ET (N = 464)
Individual Characteristics			
Age	35.67	33.31	32.41
Number of children in the household	1.22	1.67	1.77
Non-Belgian nationality	0.12	0.12	0.09
No diploma[a]	0.38	0.25	0.23
Lower secondary[a]	0.27	0.31	0.31
Higher secondary[a]	0.29	0.36	0.31
College[a]	0.06	0.08	0.08
Labour Market History in March 1997			
Unemployment duration[b]	45.74	56.01	51.88
Number of unemployment spells[c]	2.62	2.85	3.31
Duration of the last spell[d]	29.92	36.04	18.81
Last spell left censored[e]	0.27	0.30	0.11
Local Unemployment Rate	0.22	0.23	0.22

C: Controls; T: Treated; ET: Excluded Treated
All these variables come from administrative registers, with the exception of the "number of children in the household", which was reported by the individuals in the survey.

[a] "No diploma": less than 9 years of education; "Lower secondary": at least 9 years of education; "Higher secondary": at least 12 years of education; "College": at least 15 years of education

[b] Cumulative number of months passed in unemployment in the period from January 1992 to March 1997.

[c] Number of uninterrupted unemployment spells in the period from January 1992 to March 1997.

[d] Duration of the last unemployment spell, ending or still ongoing in March 1997.

[e] Binary Variable, indicating whether the last unemployment spell started prior to January 1992.

[f] Sub-regional unemployment rate for women as reported by the National Administration of UI (ONEM/RVA).

E Dehejia and Wahba (2002) Specification Test

1. Estimate the propensity score by a Logit model, with a linear specification of the covariates (no higher order terms or interactions)
2. Split the sample into equal-sized intervals based on the estimated propensity score
3. Test if the mean propensity score is not different for treated and controls within each interval
4. If (3) is rejected for one or more interval, split the interval(s) into halves and test again
5. Continue until condition (3) is satisfied
6. Test within each interval if the distribution of each covariate is not different among treated and controls
7. If there are significant differences, add some higher order terms and/or interactions, and repeat steps (1) to (6)

F Standard Errors for the Sensitivity Analysis

In this appendix, we will provide the formulas used to compute the standard errors of the average treatment effects of the simple and IV matching estimators from Sections 4.2.2 and 4.2.1. It is convenient to first consider the ATT's for each cohort of women separately according to the month in which benefits exhaust, i.e. between March and June 1997. The total average treatment on the treated at k (Δ_k), can be written as a weighted average of cohort-specific ATT's (Δ_{kt}):

$$ATT_k \equiv \Delta_k = \sum_{t=1}^{4} \frac{N_{W_t}}{N_W} \Delta_t$$

where $W \in \{T, Z_1\}$, T is the set of treated observations, N_{W_t} is the number of treated ($W = T$) or eligible ($W = Z_1$) observations at calendar time t[47], and $N_W = \sum_{t=1}^{4} N_{W_t}$. In the sequel we denote the simple and IV matching estimators respectively by superscript "S" and "IV". The simple nearest-neighbor matching estimator, Δ_{kt}^S is defined as follows:

$$\Delta_{kt}^S = \frac{1}{N_{T_t}} \left[\sum_{i \in T_t} Y_{it} - Y_{j(i)t} \right]$$

where T_t is the set of treated women of cohort t and $j(i) = argmin_{j \in C}[\hat{P}(x_i) - \hat{P}(x_j)]^2$, where C is the set of all control units. Now define ω_{ij} such that $\omega_{ij} = 1$ if $j = j(i)$ and $\omega_{ij} = 0$ if $j \neq j(i)$. If $w_{jt} \equiv \sum_{i \in T_t} \omega_{ij}$ then we can rewrite Δ_t^S alternatively as follows:

$$\Delta_{kt}^S = \frac{1}{N_{T_t}} \left[\sum_{i \in T_t} Y_{it} - \sum_{j \in C} w_{jt} Y_{jt} \right]$$

Consequently, the variance of Δ_{tk}^S is given by

$$\text{var}\left(\Delta_{kt}^S\right) = \left[N_{T_t} \text{var}\left(Y_{it}\right) + \sum_{j \in C} w_{jt}^2 \text{var}\left(Y_{jt}\right) \right] \Big/ \left(N_{T_t}\right)^2$$

The IV nearest-neighbor matching estimator is given in Equation 7. Using the corresponding definition for w_{jt}, we can write Δ_{kt}^{IV} equivalently as follows:

$$\Delta_{kt}^{IV} = \frac{1}{N_{Z_{1t}}} \left[\sum_{i \in Z_{1t}} \frac{Y_{ti}}{\hat{P}_1(x_i)} - \sum_{j \in Z_0} w_{jt} \frac{Y_{jt}}{\hat{P}_1(x_j)} \right]$$

and its variance by

$$\text{var}\left(\Delta_{kt}^{IV}\right) = \left[N_{Z_{1t}} \text{var}\left(Y_{it}\right) \Big/ \left(\hat{\text{P}}_1\left(x_i\right)\right)^2 + \sum_{j \in Z_0} w_{jt}^2 \text{var}\left(Y_{jt}\right) \Big/ \left(\hat{\text{P}}_1\left(x_j\right)\right)^2 \right] \Big/ \left(N_{Z_{1t}}\right)^2$$

Using the appropriate variance for the cohort-specific treatment effect, the variance of the total treatment effect is then computed as

[47] For each given k, there are four calendar times t, each corresponding to one of the time at which benefits exhausted plus k.

$$\mathrm{var}(\Delta_k^S) = \sum_t \left(\frac{N_{T_t}}{N_T}\right)^2 \mathrm{var}\left(\Delta_{kt}^S\right) + \frac{2}{(N_T)^2} \sum_{j \in C} \sum_t \sum_{s \neq t} w_{jt}\omega_{js}\mathrm{cov}(Y_{jt}, Y_{js})$$

and

$$\mathrm{var}(\Delta_k^{IV}) = \sum_t \left(\frac{N_{Z_1 t}}{N_{Z_1}}\right)^2 \mathrm{var}\left(\Delta_{kt}^{IV}\right) + \frac{2}{(N_{Z_1})^2} \sum_{j \in Z_0} \sum_t \sum_{s \neq t} w_{jt}\omega_{js} \frac{\mathrm{cov}(Y_{jt}, Y_{js})}{\hat{P}_1(x_j)^2}$$

References

Abbring, J. H. and G. J. van den Berg [2003], "The Non-Parametric Identification of Treatment Effects in Duration Models," *Econometrica*, 71, 1491–1517.

Abbring, J. H., G. J. van den Berg, and J. C. van Ours [2000], "The Effect of Unemployment Insurance Sanctions on the Transition Rate from Unemployment to Employment," Unpublished Manuscript, Tinbergen Institute, Amsterdam.

Ashenfelter, O. C. [1978], "Estimating the Effect of Training Programs on Earnings with Longitudinal Data," *The Review of Economics and Statistics*, 6 (1), 47–57.

Athey, S. and G. Imbens [2002], "Identification and Inference in Nonlinear Difference-In-Differences Models," Working Paper T0280, NBER.

Atkinson, A. B. and J. Micklewright [1991], "Unemployment Compensation and Labor Market Transitions: A Critical Review," *Journal of Economic Literature*, XXIX, 1679–1727.

Blundell, R. and M. Costa-Dias [2002], "Alternative Approaches to Evaluation in Empirical Microeconomics," *Portuguese Economic Journal*, 1 (2), 91–115.

Blundell, R., M. Costa-Dias, C. Meghir, and J. Van Reenen [2001], "Evaluating the Employment Impact of a Mandatory Job Search Assistance Program," Working Paper W01/20, Institute for Fiscal Studies, London.

Bover, O., M. Arellano, and S. Bentolila [2002], "Unemployment Duration, Benefit Duration and the Business Cycle," *Economic Journal*, 112, 223–265.

Bratberg, E. and K. Vaage [2000], "Spell Durations with Long Unemployment Insurance Periods," *Labour Economics*, 7 (2), 153–180.

Card, D. and P. B. Levine [2000], "Extended Benefits and the Duration of UI Spells: Evidence from the New Jersey Extended Benefit Program," *Journal of Public Economics*, 78 (1–2), 107–138.

Carling, K., P.-A. Edin, A. Harkman, and B. Holmlund [1996], "Unemployment Duration, Unemployment Benefits and Labor Market Programs in Sweden," *Journal of Public Economics*, 59, 313–334.

Carling, K., B. Holmlund, and A. Vejsiu [2001], "Do Benefit Cuts Boost Job Findings? Swedish Evidence from the 1990's," *Economic Journal*, 111, 766–790.

De Lathouwer, L., K. Bogaerts, and S. Perelman [2000], *Exclusion de l'assurance chômage et réinsertion sur le marché du travail*, Services fédéraux des affaires scientifiques techniques et culturelles, Bruxelles.

De Lathouwer, L., K. Bogaerts, and K. Van den Bosch [2003], "Een evaluatie van schorsing artikel 80 vanuit herintrede- en armoedsperpectief," in L. De Lathouwer, B. Cockx, K. Bogaerts, J. Ries, and K. Van den

Bosch (eds.), *L'impact de la suspension article 80 dans l'assurance–chômage sur la réinsertion et la pauvreté*, Academia Press, Gent.

Dehejia, R. H. and S. Wahba [1999], "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, 94 (448), 1053–1062.

Dehejia, R. H. and S. Wahba [2002], "Propensity Score Matching Methods for Non-Experimental Causal Studies," *The Review of Economics and Statistics*, 84, 151–161.

Devine, T. J. and N. M. Kiefer [1991], *Empirical Labor Economics. The Search Approach*, Oxford University Press, New York.

Dormont, B., D. Fougère, and A. Prieto [2001], "L'effet de l'allocation unique dégressive sur la reprise de l'emploi," *Economie et Statistique*, 343 (3), 3–28.

Eichler, M. and M. Lechner [2002], "An Evaluation of Public Employment Programmes in the East German State of Sachsen-Anhalt," *Labour Economics*, 9, 143–186.

Fredriksson, P. and B. Holmlund [2003a], "Improving Incentives in Unemployment Insurance: A Review of Recent Research," Working Paper 2003:10, Department of Economics, Uppsala University.

Fredriksson, P. and B. Holmlund [2003b], "Optimal Unemployment Insurance Design: Time Limits, Monitoring, or Workfare?" Working Paper 2003:17, Department of Economics, Uppsala University.

Hahn, J., P. Todd, and W. van der Klauw [2001], "Identification and Estimation of Treatment Effects by Regression Discontinuity Design," *Econometrica*, 63 (3), 201–209.

Ham, J. C. and S. A. Rea [1987], "Unemployment Insurance and Male Unemployment Duration in Canada," *Journal of Labor Economics*, 5, 325–353.

Heckman, J. J. and V. J. Hotz [1989], "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training," *Journal of the American Statistical Association*, 84 (408), 862–880.

Heckman, J. J., H. Ichimura, and P. E. Todd [1997], "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies*, 64, 605–654.

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

Heckman, J. J., R. J. Lalonde, and J. A. Smith [1999], "The Economics and Econometrics of Active Labor Market Programs," in O. C. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 3A, chap. 31, North-Holland, Amsterdam, 1277–1366.

Holmlund, B. [1998], "Unemployment Insurance in Theory and Practice," *Scandinavian Journal of Economics*, 100 (1), 113–141.

Hunt, J. [1995], "The Effect of Unemployment Compensation on Unemployment Duration in Germany," *Journal of Labor Economics*, 13 (1), 88–120.

Ichimura, H. and C. Taber [2001], "Propensity Score Matching with Instrumental Variables," *American Economic Review*, 91 (2), 119–124.

Imbens, G. W. [2003], "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review," Working Paper T0294, NBER.

Imbens, G. W. and J. D. Angrist [1994], "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62 (2), 467–475.

Jenkins, S. P. and C. Garcia-Serrano [2000], "Re-employment Probabilities for Spanish Men: What Role Does the Unemployment Benefit System play ?" Working Paper 2000–17, ISER, University of Essex.

Jensen, P., M. Rosholm, and M. Svarer [2003], "The Response of Youth Unemployment to Benefits, Incentives, and Sanctions," *European Journal of Political Economy*, 19 (2), 301–316.

Joutard, X. and M. Ruggiero [1996], "Changements de régime d'indemnisation et transitions vers l'emploi," *Revue Economique*, 47 (1), 143–166.

Katz, L. F. and B. D. Meyer [1990], "The Impact of Potential Benefit Duration of Unemployment Benefits on the Duration of Unemployment," *Journal of Public Economics*, 41, 45–72.

Lalive, R., J. C. van Ours, and J. Zweimüller [2002], "The Effect of Benefit Sanctions on the Duration of Unemployment," Discussion Paper 469, IZA, Bonn.

Lalive, R. and J. Zweimüller [2002], "Benefit Entitlement and Unemployment Duration: The Role of Policy Endogeneity," Discussion Paper 492, IZA, Bonn.

Layard, R., S. Nickell, and R. Jackman [1991], *Unemployment: Macroeconomic Performance and the Labour Market*, Oxford University Press, Oxford.

Lechner, M. [2001], "A Note on the Common Support Problem in Applied Evaluation Studies," Discussion Paper 2001–01, Department of Economics, University of St. Gallen.

Lindeboom, M. and J. Theeuwes [1993], "Search, Benefits and Entitlement," *Economica*, 60, 327–346.

Moffit, R. [1985], "Unemployment Insurance and the Distribution of Unemployment Spells," *Journal of Econometrics*, 28 (1), 85–101.

Mortensen, D. T. [1977], "Unemployment Insurance and Job Search Decisions," *Industrial and Labor Relations Review*, 30, 505–517.

Mortensen, D. T. [1990], "A Structural Model of Unemployment Insurance Benefit Effects on the Incidence and Duration of Unemployment," in Y. Weiss and G. Fishelson (eds.), *Advances in the Theory and Measurement of Unemployment*, Macmillan, Hampshire.

Pedersen, P. J. and N. Westergård–Nielsen [1998], "Unemployment: What Do we Know from Longitudinal Data?" in T. Lange (ed.), *Unemployment in Theory and Practice*, Edward Elgar, Cheltenham.

Roed, K. and T. Zhang [2003], "Does Unemployment Compensation Affect Unemployment Duration?" *Economic Journal*, 113 (484), 190–206.

Rosenbaum, P. R. and D. B. Rubin [1983], "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70 (1), 41–55.

Rosenbaum, P. R. and D. B. Rubin [1984], "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score," *Journal of the American Statistical Association*, 79 (387), 516–524.

Silverman, B. [1986], *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Smith, J. A. and P. E. Todd [2004], "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators," *Journal of Econometrics*, forthcoming.

Thoursie, A. [1998], "Effects of Renewable Benefit Periods on the Exit Rate from Unemployment," in A. Thoursie (ed.), *Studies on Unemployment Duration and on the Gender Wage Gap*, no. 35 in Dissertation Series, Swedish Institute for Social Research, Stockholm.

van den Berg, G. J. [1990], "Nonstationarity in Job Search Theory," *Review of Economic Studies*, 57, 255–277.

Winter-Ebmer, R. [1998], "Potential Unemployment Benefit Duration and Spell Length: Lessons from a Quasi-Experiment in Austria," *Oxford Bulletin of Economics and Statistics*, 60 (1), 33–45.

Wooldridge, J. M. [2002], *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge.

Wurzel, E. [1990], "Staggered Entry and Uneployment Durations: an Application to German Data," in J. Hartog, G. Ridder, and J. Theeuwes (eds.), *Panel Data and Labor Market Studies*, North–Holland, Amsterdam.