UNEMPLOYMENT DURATION, INCENTIVES AND INSTITUTIONS. A MICRO-ECONOMETRIC ANALYSIS BASED ON SCANDINAVIAN DATA

by

Knut Røed, Peter Jensen, Anna Thoursie
Unemployment Duration, Incentives and Institutions
- A Micro-Econometric Analysis Based on Scandinavian Data*

Knut Røed
The Ragnar Frisch Centre for Economic Research, Oslo, Norway
knut.roed@frisch.uio.no

Peter Jensen
Department of Economics, The Aarhus School of Business, Denmark
pje@asb.dk

Anna Thoursie
Swedish Institute for Social Research, Stockholm, Sweden and the
Swedish Trade Union Confederation, Stockholm, Sweden
Anna.Thoursie@lo.se

Abstract

Based on a combined register database for Norwegian and Swedish unemployment spells, we use the ‘between-countries-variation’ in the unemployment insurance systems to identify causal effects. The elasticity of the job hazard rate with respect to the benefit replacement ratio is around -1.0 in Norway and -0.5 in Sweden. The limited benefit duration period in Sweden has a large positive impact on the hazard rate, despite generous renewal options through participation in labour market programs. Compulsory program participation seems to operate as a ‘stick’, rather than a ‘carrot’, and is therefore an efficient tool for counteracting moral hazard problems in the benefit system.

Keywords: Unemployment spells, unemployment compensation, non-parametric duration analysis.

JEL Classification: C41, J64.

* This research is financed by the Nordic Council of Ministers and the Joint Committee of the Nordic Research Councils for the Social Sciences. We wish to thank Pål Børing, Erik Hernæs and Tor Eriksson for valuable comments, Øystein Jørgensen and Anders Wellman for skilful preparation of the data and Simen Gaure for valuable programming assistance. Correspondence to: Knut Røed, The Ragnar Frisch Centre for Economic Research, Gaustadalleen 21, 0349 Oslo, Norway. E-mail: knut.roed@frisch.uio.no.
1 Introduction

The question of how economic incentives embedded in the unemployment compensation system affect transitions out of unemployment has received considerable attention in the literature. Surveys are provided by Danziger et al (1981), Devine and Kiefer (1991), Atkinson and Micklewright (1991), Layard et al (1991), and more recently by Holmlund (1998) and Pedersen and Westergård-Nielsen (1998). While the theoretical literature (Mortensen, 1977; 1990; van den Berg, 1990b) clearly predicts that a higher level of unemployment insurance (UI) benefits reduces the transition rate from unemployment to employment (at least for newly unemployed benefit claimants), the lessons arising from the empirical literature are more mixed. Empirical results for the United States and the United Kingdom (Lancaster, 1979; Moffitt, 1985; Narendranathan et al, 1985; Katz and Meyer, 1990; Meyer, 1990) tends rather unanimously to favour significant effects associated with the UI benefit level. And the conclusion reached by Lancaster (1979, p. 956) more than 20 years ago that an elasticity of the escape rate from unemployment with respect to the benefit level in the order of -0.6 could be regarded as ‘established beyond reasonable doubt’, still seems rather representative. Evidence from central parts of Europe (Hujer and Schneider, 1989; Groot, 1990; van den Berg, 1990a; Steiner, 1990; Wurzel, 1990; Lindeboom and Theeuwes, 1993; Hunt, 1995; van den Berg et al, 1998) is more ambiguous, and significant benefit level effects are often hard to establish at all. There has been a number of Scandinavian contributions to the literature in recent years; all of them emanating from the fast growing supply of register based data sets (Jensen and Westergård-Nielsen, 1990; Hernæs and Strøm, 1996; Carling et al, 1996; Carling et al, 2001; Røed and Zhang, 2003). They tend to support the Anglo-American finding of a
significant disincentive effect associated with the UI benefit level, although the size of the effect is far from established.

There are doubts regarding the validity of all these results, however. The empirical problem is that there is virtually no truly independent variation in individual UI benefit entitlements that can be used to identify causal effects. Variations in UI benefits, as well as in replacement ratios, are typically strongly correlated to variations in previous income, which again are correlated to unobserved characteristics that themselves affect the escape rate from unemployment. One approach that has been adopted by researchers in the United States is to take advantage of institutional differences between states (Moffitt, 1985; Katz and Meyer, 1990). Katz and Meyer (1990), for example, use the UI benefit and the previous income as explanatory variables in a hazard rate model, and then rely on differences in the functional non-linear relationship between income and benefits across states to identify the causal effect of interest. Potential differences in hazard rates between states that are not related to the UI benefit system are controlled for by means of state dummy variables. The main problem with this approach is that the identification of causality rests on an unjustified and rather restrictive (log) linearity assumption regarding the ‘effect’ of previous (or expected) income. If the income affects the hazard rate in a non-linear fashion (either through direct causality or through linkages to unobserved characteristics), any non-linear function of previous income may pick this effect up, including the UI benefit level. Hence, the estimated effect of UI benefits may still be spurious. Moreover, there may be differences in the hazard rate pattern across states that are not fully accounted for by dummy variables. Another approach that has been used by some researchers to solve the problem of spurious correlation is to take advantage of
structural reforms (e.g. changes in the compensation level) that affect some, but not all unemployed persons, and apply the difference-in-difference methodology (Meyer, 1989; Hunt, 1995; Winter-Ebmer, 1998; Carling et al, 2001). This approach may also be problematic, however, as it rests on the often questionable assumption that labour market opportunities do not develop differently for the ‘treatment’ and the ‘control’ groups. Røed and Zhang (2003) take advantage of some particularities of the Norwegian UI benefit system that produce an element of random (from the individuals’ viewpoint) variation in replacement ratios. They find strong disincentive effects, with benefit elasticities around –0.7 on average. But ‘randomised’ variation in benefits is indeed a rarity; hence the generality and significance of these results remains to be verified.

Scandinavian policy makers have long realised that their rather generous income support systems may entail a moral hazard problem. But, rather than lowering the level of benefits, the main strategy has been to accompany the income support measures with ‘carrots and sticks’ aimed at counteracting the disincentive effects. This policy has in particular been associated with strict eligibility criteria regarding willingness to accept available work (OECD, 2000) and limited duration of benefits. But, although the literature is more unanimous in terms of identifying significant effects of benefit exhaustion and sanctions (Meyer, 1990; Hunt, 1995; Carling et al, 1996; Winter-Ebmer, 1998; Thoursie, 1998; Bratberg and Vaage, 2000; Card and Levine, 2000; Røed and Zhang, 2003), there are again methodological difficulties that cast doubt on the results. First, the transition rate may change over spell duration for a number of reasons, such as discouragement, depreciation of human capital, statistical discrimination, information gathering etc.; hence it is difficult to isolate the effect of benefit exhaustion. Second, the concepts
of ‘limited benefit duration’ and ‘sanctions’ mean different things in different countries. While for example the US version of these concepts truly implies economic hardship for the people concerned, the Scandinavian versions basically imply that benefits are replaced by program participation income or by means tested cash transfers. This suggests that empirical results regarding threats of benefit termination are not easily transferable across different countries.

The aim of the present paper is to use the variation in incentive-structures between unemployed persons in Norway and Sweden in order to identify causal effects of the benefit system on the job transition rates. The basic idea is to identify the effect of UI benefits by ‘purifying’ the variation that occurs solely because people happen to live in countries with different UI institutions. Norway and Sweden are in many respects extremely similar: They were united until 1905, they are culturally more or less indistinguishable, they have (almost) the same language and they have since 1954 adhered to a common Nordic labour market. They have both very high labour market participation rates, particularly among women. They share many of the same institutions, with relatively large welfare states, highly regulated labour markets and relatively co-ordinated wage setting. However, their UI benefit systems differ quite sharply in two respects: The profile of replacement ratios and the maximum duration of benefit eligibility. Although the replacement ratio declines with income in both countries, the exact functional relationship differs. While the Swedish system provides the highest replacement ratios for low-income workers, the Norwegian system is the most generous for high-income workers. Hence, to the extent that the replacement ratio is an important determinant of the transition rate from unemployment to employment, one would expect to find that the ratio
of the transition rates for low- to high-income workers is lower in Sweden than in Norway. There are also substantial differences in the maximum duration by which benefits can be maintained without ‘activity’ requirements. While benefits can be obtained for 156 weeks in Norway they are limited to 60 weeks in Sweden. To the extent that these constraints affect behaviour, one would expect to find traces of the different systems in the pattern of job transition rates, e.g. in the form of rising hazard rates in the period just prior to benefit exhaustion.

In the present paper, we provide compelling evidence that the institutional differences between the Norwegian and the Swedish UI benefit regimes indeed have the predicted consequences referred to above. The two different systems are clearly mirrored in the data in a way that convincingly establishes the link between economic incentives and job search behaviour. The paper proceeds as follows: The next section describes our data, with emphasis on the structural differences between the two countries. Section 3 presents the econometric model that we use to extract the causal parameters of interest. Section 4 presents the main results, and Section 5 concludes.

2 Data

We have gathered complete register data for unemployment spells in Norway and Sweden during 1999 and 2000. From these databases, we have selected entrants into ordinary full-time unemployment aged 25-54 who were eligible for UI benefits (and who would have been so in both countries) and who were employed prior to the unemployment spell. Their subsequent unemployment status was then observed by the end of each calendar month during the two-year data period. For each person, we have information about previous income, UI benefits, age, gender, educational attainment, job practice, nationality
and region. Despite the differences in registration routines between the two countries, we consider the Norwegian and Swedish data to be highly comparable. Descriptive statistics are provided in Table 1. In total, there are around 1.2 million monthly observations divided into 327,000 spells shared between 250,000 individuals. Swedish spells were on average much shorter than Norwegian spells. On the other hand, while two thirds of the completed Norwegian spells ended directly in a transition to employment\(^1\), this was the case for only 42 per cent of the Swedish spells. As much as 35 per cent of the completed Swedish spells ended in a transition to a labour market program, while this was the case for only 17 per cent of the Norwegian spells. Hence, it seems that Swedish labour market policy is the most active, in the sense that the job seekers are spurred into some kind of activity if a job is not obtained fairly quickly. The average monthly transition rate from unemployment to employment was around 11 per cent in Sweden and 10 per cent in Norway. It may also be noted that the ratio of job transition rates for the lowest versus the highest decile in the country specific wage distribution is much lower in Sweden (0.69) than in Norway (0.92). This is a first rough indication that the differences in the replacement ratio profiles between the two countries do have the expected effects on the relative exit rates.

<table>
<thead>
<tr>
<th></th>
<th>Norway</th>
<th>Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of individuals</td>
<td>31,151</td>
<td>218,545</td>
</tr>
<tr>
<td>Number of spells</td>
<td>35,969</td>
<td>291,229</td>
</tr>
<tr>
<td>Number of monthly observations</td>
<td>188,406</td>
<td>1,080,040</td>
</tr>
</tbody>
</table>

\(^1\) We are not able to identify the entry into jobs directly for both countries; hence we assume that persons who leave the unemployment register completely have obtained jobs. Given that we focus on a group of persons with strong pecuniary incentives to register in the absence of a job transition, we consider this assumption to be reasonable.
Table 1
Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Norway</th>
<th>Sweden</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome of spells (per cent)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transition to a job (out of the register)</td>
<td>48.74</td>
<td>37.05</td>
</tr>
<tr>
<td>Transition to a labour market program</td>
<td>13.09</td>
<td>31.32</td>
</tr>
<tr>
<td>Other transition (loss of benefits or disability)</td>
<td>14.36</td>
<td>19.87</td>
</tr>
<tr>
<td>Censored at the end of the observation window</td>
<td>23.81</td>
<td>11.76</td>
</tr>
<tr>
<td><strong>Average monthly transition probability to a job (per cent)</strong></td>
<td>9.86</td>
<td>11.10</td>
</tr>
<tr>
<td><strong>Transition to a job for the poorest decile relative to the richest decile</strong></td>
<td>0.92</td>
<td>0.69</td>
</tr>
<tr>
<td><strong>Fraction women (per cent of individuals)</strong></td>
<td>44.37</td>
<td>46.39</td>
</tr>
<tr>
<td><strong>Age (average taken over individuals)</strong></td>
<td>36.33</td>
<td>36.00</td>
</tr>
<tr>
<td><strong>Educational attainment (per cent of individuals)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only compulsory school</td>
<td>23.36</td>
<td>21.56</td>
</tr>
<tr>
<td>Incomplete secondary school</td>
<td>17.21</td>
<td>6.83</td>
</tr>
<tr>
<td>Completed secondary school</td>
<td>39.16</td>
<td>54.77</td>
</tr>
<tr>
<td>1-3 year College/University</td>
<td>15.95</td>
<td>10.09</td>
</tr>
<tr>
<td>More than 3 year College/University</td>
<td>4.32</td>
<td>6.76</td>
</tr>
<tr>
<td><strong>Citizenship (per cent of individuals)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>National</td>
<td>90.76</td>
<td>91.62</td>
</tr>
<tr>
<td>Other Nordic</td>
<td>1.44</td>
<td>2.24</td>
</tr>
<tr>
<td>Other OECD</td>
<td>3.56</td>
<td>3.01</td>
</tr>
<tr>
<td>Non-OECD</td>
<td>4.24</td>
<td>3.13</td>
</tr>
</tbody>
</table>

Figure 1 describes the distribution of replacement ratios and previous (daily) incomes in the two countries. The difference in replacement ratio profiles is illustrated in panel (a). It is clearly seen that Sweden has higher replacement ratios than Norway for low incomes and lower replacement ratios for high incomes (with incomes measured in a common currency). However, panel (b) indicates that there is also a more equal distribution in previous incomes among unemployed persons in Sweden than in Norway and that the income level is also somewhat higher in Norway. As a result, it can be seen from panel (c) that the replacement ratio is higher in Sweden than in Norway at virtually all percentiles of each country’s income distribution.
Figure 1. Distribution of previous income and replacement ratios in Norway and Sweden, based on the lower bounds of the percentiles in each country’s income distribution

While benefits can be maintained almost indefinitely in Norway (156 weeks, in many cases followed by a somewhat smaller cash-transfer), the Swedish benefit period is limited to 60 weeks. Until 2001 (i.e. during the whole period covered by our data), benefit entitlements in Sweden could be renewed through participation in labour market programs. It is important to bear in mind that in both countries, the duration constraint is a very ‘soft’ one, in the sense that paid labour market programs and/or means tested Social Security benefits are available for persons without benefit entitlements. However, some form of ‘activation’ is usually required when the benefit entitlements are exhausted.

The period covered by our data was a period of sturdy labour demand in both countries, and in Sweden it was also a period of strong recovery. The standardised unem-
ployment rates in 1999 and 2000 were 3.2 and 3.5 per cent in Norway, and 7.2 and 5.9 per cent in Sweden.

3 Econometric Approach

The analysis in this paper focuses particularly on two aspects of the benefit systems, the *replacement ratio* and *benefit exhaustion*. The basic idea is to take advantage of the between-countries variation in these institutional factors in order to eliminate bias in causal parameters arising from correlation between the variables of interest and unobserved individual characteristics. Hence, we identify the effect of the replacement ratio through the variation in replacement ratios between individuals with similar incomes, but with different replacement ratios because they happen to live in countries with different rules. The spurious part, reflecting that persons with different incomes also may have different unobserved characteristics, is eliminated through the inclusion of both countries’ hypothetically calculated replacement ratios as explanatory variables. In this paper, we choose to represent the UI benefit in the form of a (log) *replacement ratio*. This implies that it is only the ratio between benefits and the expected wage that affects behaviour, and not their respective levels. This greatly facilitates compatibility between the two countries, but the underlying homogeneity assumption is also in line with influential theoretical predictions arising from dynamic search theory (see e.g. Mortensen, 1990, pp 68-69). The expected wage is assumed to equal the previously observed wage.

Let \( \theta_i(t, d) \) be the hazard rate by which individual \( i \) transits to employment at calendar time \( t \), when spell duration is \( d \). As we observe labour market status by the end of each month, we set up the model in terms of discrete (grouped) hazard rates. Let \( \tilde{r}_i \) be the
calendar time at which individual \( i \) entered the state of unemployment. The grouped hazard, i.e. the probability of exiting during duration month \( d \), given that no exit occurred before that, is given as

\[
h_{id} = 1 - \exp\left( -\int_{d-i}^d \theta_i(t_i + u) du \right). \tag{1}
\]

Assume that the hazard rate is proportional in factors depending on observed covariates, unobserved covariates, calendar-time and process-time (spell duration). Assume also for simplicity that time-varying covariates only vary between (and not within) observation months, and that spell duration effects are constant within each month. These assumptions imply that we can write the integrated hazard rate in (1) in terms of integrals taken over the calendar-month component only, multiplied by factors of proportionality reflecting spell duration and effects of observed- and unobserved heterogeneity, respectively. Let \( x_i \) be a vector of individual control variables, such as age, gender, educational attainment etc., and let \( r_i \) be the log of the replacement ratio applying for individual \( i \).

Furthermore, let \( \sigma_{jt} \) and \( \lambda_{jd} \) be measures of the calendar time- and spell duration effects respectively, in each country \( j \). Imposing exponential link functions between individual characteristics and the hazard rates, we then have that the probability that a person \( i \) exits unemployment during calendar month \( t \), when the spell duration is \( d \) months is given as:

\[
h_{id} = h(t, d, x_i, r_i, v_i) = 1 - \exp\left( -\exp\left( x_i \beta_j + \kappa r_i + \sigma_{jt} + \lambda_{jd} + v_i \right) \right), \tag{2}
\]

where \( v_i \) is a scalar measure of unobserved heterogeneity. Note that the calendar time effects are not restricted to remain constant within each month, and that \( \sigma_{jt} \) is the (log of the) integral taken over the calendar time component corresponding to the month in question. Indeed, it is likely that the transition rates vary substantially over the calendar
month, e.g. with much higher transition rates on Mondays than on other weekdays. With suitable trading day- and seasonal adjustments, the $\sigma_{j\mu}$ can be given a business-cycle interpretation (Røed, 2001). The parameters of interest in the present paper are the $\kappa_{j}'s$, i.e. the elasticities of the hazard rates with respect to the replacement ratio, and the $\lambda_{ijd}'s$, i.e. the spell duration effects. The problem with identifying these parameters is that both the replacement ratios and the spell durations are likely to be highly correlated to unobserved heterogeneity.

Consider first the scope for identifying the replacement ratio effects. The correlation between replacement ratios and unobserved heterogeneity arises because replacement ratios depend on previous income, whereas previous income (conditioned on other explanatory variables) is likely to be correlated to unobserved characteristics such as motivation and working spirit (represented by the term $v_i$). A direct estimation of (2) in disregard of unobserved heterogeneity thus produces a coefficient attached to the replacement ratio that is a mixture of the causal effect embedded in the parameter $\kappa_{j}$ and the correlation between the replacement ratio and the unobserved covariate $v_i$. In order to remove this source of bias, we assume that the correlation structure between the replacement ratios calculated according to the different country-specific rules and unobserved heterogeneity is common across the two countries, i.e.

$$v_i = \alpha_1 r_{IN} + \alpha_2 r_{IS} + \varepsilon_i,$$  \hspace{1cm} (3)

where $r_{IN}$ is the replacement ratio applying for individual $i$ if that individual register in Norway, and $r_{IS}$ is the ratio applying if registering in Sweden. Hence, a consistent estimate of $\kappa_{j}$ can be obtained by including both hypothetical replacement ratios as ex-
planatory variables for the transition rates, and assuming that they have the same ‘effect’ in both countries. The intuition is that while the replacement ratio calculated according to one country’s specific rules can have causal effects only in that particular country, its spurious effects applies to the other country as well. The hypothetical replacement ratios are calculated on the basis of the unemployed workers position in the distribution of previous incomes among the unemployed (see Figure 1, panel c). For example, a Swedish unemployed person with a previous income that places him in the p’th decile in the income distribution among the Swedish unemployed, is equipped with a hypothetical Norwegian replacement ratio equal to that faced by a Norwegian unemployed belonging to the p’th decile in the Norwegian income distribution. Inserting (3) in (2) yields

\[ h_{itd} = h(i, d, x_{it}, \varepsilon_i) = 1 - \exp\left(-\exp\left(x_{it} \beta_j + \kappa j_i + \lambda_{jd} + \alpha_1 r_{it} + \alpha_2 r_{is} + \varepsilon_i\right)\right), \]  

(4)

where \( \varepsilon_i \) is now orthogonal to \( r_i \).

Consider then the identification of the spell duration effects, embedded in \( \lambda_{jd} \). A common problem in duration analysis is that the duration patterns arising from unobserved heterogeneity and structural duration dependence are indistinguishable from data alone. There are two potential sources of non-parametric identification. The first is repeated spells by the same individuals (see van den Berg, 2001). But this source of identification requires that there is no dependence across spells apart from that arising from the person-specific ‘fixed effect’ \( \varepsilon_i \) (or that this dependence is known). This is a questionable assumption insofar as there are positive or negative duration effects within each spell, particularly if the spells are close in time (an apparently new spell may in some senses be a continuation of the previous one). Moreover, the actual usage of this source of identification may in our case introduce more serious selection problems than it solves, since the
probability of having more than one spell during a given two year period obviously is strongly correlated to the length of the first spell. The second source of identification builds on variation in lagged explanatory variables. Its basic idea is that the conditional expectation of unobserved heterogeneity (i.e. conditioned on observed characteristics and spell duration) depends on exit rates experienced earlier in the (same) spell, while true duration dependence does not (van den Berg and van Ours, 1994; 1996). The higher the exit rates have been earlier in a spell, the more selection has taken place at any given duration, and the lower is the expected value of the unobserved covariate $\varepsilon_i$. In the present context, the existence of multiple cohorts ensures that persons with exactly the same spell durations have been subject to different labour market conditions earlier in the spell, and hence have been exposed to different selection forces. This variation in lagged explanatory variables is substantial and makes it possible to identify the underlying structural duration dependence based on the information content in the data alone, i.e. without any arbitrary parametric assumptions at all. Identification is obtained even without the assumption of a Mixed Proportional Hazard model (Brinch, 2000).

In order to avoid unnecessary parametric restrictions, we assume that the unobserved variables $\varepsilon_i$ is discretely distributed (Lindsay, 1983), with the number of mass-points chosen by adding points until it is no longer possible to increase the likelihood function (Heckman and Singer, 1984). The formulation of the likelihood function depends on the sources used for identification of unobserved heterogeneity. When both sources are used (repeated spells and lagged explanatory variables), unobserved characteristics are assumed fixed at the individual level across different spells. Let $B_i$ be the number of spells experienced by individual $i$ during the whole observation period. Each
spell either ends in a transition, or it is censored. Censoring occurs when persons exit to labour market programs, when benefits are terminated, and at the end of the observation period. Let $y_{ib}=1$ if spell number $b$ of individual $i$ ends in a transition (non-censored), and zero otherwise, let $d_{ib}$ be the duration of that spell. Let $W$ be the number of mass points in the distribution of unobservables and let $p_w$ be the probability that the unobserved covariate obtains the value $\varepsilon_w$. The likelihood function in terms of observations of $(d_{ib}, y_{ib}, t, x_{it})$ is then given as

$$L_I = \prod_{i=1}^{N_I} \sum_{w=1}^{W} \sum_{b=1}^{B_i} p_w \prod_{s=1}^{d_{ib}-1} (1 - h(t-s, d_{ib} - s, x_{it-s}, \varepsilon_w)) \prod_{s=y_{ib}}^{d_{ib}-1} h(t-s, d_{ib} - s, x_{it-s}, \varepsilon_w) \prod_{w} p_w = 1,$$

where $N_I$ is the number of individuals in the dataset. The likelihood (5) is maximised with respect to the model parameters in (4) and with respect to the parameters entering the discrete distribution of unobserved heterogeneity $(W, p_w, \varepsilon_w)$. When only the second source of identification (lagged explanatory variables) is trusted, each spell in the dataset is treated as a different individual (hence $B_i = 1$ for all ‘individuals’), and the likelihood function in terms of observations of $(d_i, y_i, t, x_{it})$ becomes

$$L_S = \prod_{i=1}^{N_S} \sum_{w=1}^{W} p_w \left( (h(t, d_i, x_{it}, \varepsilon_w))^{y_{i}} \prod_{s=y_{i}}^{d_i-1} (1 - h(t-s, d_i - s, x_{it-s}, \varepsilon_w)) \right) \prod_{w} p_w = 1,$$

where $N_S$ is the number of spells in the dataset ($N_S > N_I$) and $i$ is the identifier of spells, rather than individuals.

Estimates based on (6) are consistent either the fixed-effect-across-spells assumption holds or not, but they are not efficient when it holds. Hence, we can use a Hausman test (Hausman, 1978) to confront the fixed-effect-across-spells assumption embedded in (5) with data.
4 Results

The model was estimated both with and without the fixed-effect-across-spells assumption. In the former case (equation 5), the model was ‘saturated’ with seven points of support in the unobserved heterogeneity distribution, while in the latter case (equation 6), it was sufficient with three points. The two models produced similar results with respect to replacement ratio elasticities and effects of benefit exhaustion (i.e. the changes in the baseline in the months just prior to exhaustion). However, the rest of the estimated duration baseline differed quite sharply, with strong positive duration dependence in both countries implied by the fixed-effect-across-spells model, and no systematic, or even negative duration dependence implied by the least restrictive model. The Hausman test strongly rejected the fixed-effect-across-spells assumption\(^2\). Hence, we conclude that the data confirmed our suspicions regarding the reliance on repeated spells to identify spell duration effects. We therefore focus on the results from the least restrictive model (equation 6), even though it does not take advantage of all available information (i.e. that some persons are observed having more than one spell), and hence cannot be fully efficient.

The estimates regarding benefit effects and most individual characteristics are presented in Table 2, while the estimates regarding age and the spell duration baselines are presented in Figures 2 and 3, respectively (we use a graphic presentation of these re-

\(^2\) The Hausman test is a parametric test; hence we conditioned on the number of mass-points in the heterogeneity distribution. The test was performed in the following way: Let \(b_1\) be the vector of common parameter estimates from the model without the fixed-effect-across-spells assumption (equation 6), let \(b_2\) be the same estimates from the model based on the fixed-effect-across-spells assumption (equation 5, with the same number of mass-points), and let \(V_1\) and \(V_2\) be the associated covariance matrixes. The Hausman test statistic was then calculated as \(H = (b_1 - b_2)'(V_1 - V_2)^{-1}(b_1 - b_2)\) and under the null hypothesis that the fixed-effect-across-spells assumption holds this statistic is chi-square distributed with the degrees of freedom corresponding to the number of common parameters. With three mass-points in each model, we obtained a Chi-Square Statistic (with 215 degrees of freedom) of around 1 million.
sults due to the large number of parameters). A point to note is that most of the control variables have very similar effects in Norway and Sweden. For example, the effects associated with educational attainment (reported in Table 2) and age (reported in Figure 2) are hardly distinguishable. We interpret this as supporting evidence regarding the comparability of the data, as well as regarding the assumption that heterogeneity affects the Norwegian and Swedish hazard rates in a similar fashion. On the other hand, there are huge differences in the spell duration pattern and in the replacement ratio elasticities.

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Selected Maximum Likelihood Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Norway</td>
</tr>
<tr>
<td></td>
<td>Estimate</td>
</tr>
<tr>
<td>Replacement ratio (causal effect)</td>
<td>-1.07**</td>
</tr>
<tr>
<td>Being a woman</td>
<td>-0.07**</td>
</tr>
<tr>
<td>Educational attainment</td>
<td></td>
</tr>
<tr>
<td>Only compulsory school</td>
<td>0</td>
</tr>
<tr>
<td>Incomplete secondary school</td>
<td>0.02</td>
</tr>
<tr>
<td>Completed secondary school</td>
<td>0.08**</td>
</tr>
<tr>
<td>1-3 year College/University</td>
<td>0.20**</td>
</tr>
<tr>
<td>More than 3 year College/University</td>
<td>0.27**</td>
</tr>
<tr>
<td>Relevant education (with respect to wanted job type)</td>
<td>0.09**</td>
</tr>
<tr>
<td>Relevant practice (with respect to wanted job type)</td>
<td>0.17**</td>
</tr>
<tr>
<td>Citizenship</td>
<td></td>
</tr>
<tr>
<td>National (Norwegian/Swedish)</td>
<td>0</td>
</tr>
<tr>
<td>Other Nordic</td>
<td>0.22**</td>
</tr>
<tr>
<td>Other European</td>
<td>-0.15**</td>
</tr>
<tr>
<td>Asian (including Turkish)</td>
<td>-0.25**</td>
</tr>
<tr>
<td>African</td>
<td>-0.26**</td>
</tr>
<tr>
<td>Latin American</td>
<td>-0.01</td>
</tr>
<tr>
<td>North American or Oceanic</td>
<td>-0.24</td>
</tr>
<tr>
<td>Common Parameters</td>
<td></td>
</tr>
<tr>
<td>Norwegian replacement ratio (spurious effect)</td>
<td>1.44**</td>
</tr>
<tr>
<td>Swedish replacement ratio (spurious effect)</td>
<td>-0.75**</td>
</tr>
<tr>
<td>Mass points in distribution of unobserved variables</td>
<td>3 points (variance of ε=0.02)</td>
</tr>
<tr>
<td>Number of parameters</td>
<td>220</td>
</tr>
<tr>
<td>Log-likelihood</td>
<td>-364670.97</td>
</tr>
</tbody>
</table>

*(**) significant at the level of 5(1) per cent.

Note: In addition to the variables reported in the table and in figures, calendar month- and region (‘fylke’ in Norway, ‘län’ in Sweden) dummies were included for both countries.
Consider first the estimates of structural duration dependence depicted in Figure 3. While there is a relatively strong pattern of negative duration dependence in Norway, the hazard rate in Sweden does not display any clear duration pattern\(^3\). It rises slightly during the first six months, and then declines during the next six months. The most conspicuous feature of the Swedish spell duration baseline, however, is that it rises by around

\(^3\) To check out the robustness of the results, we have also estimated the spell duration baselines in completely separate models for the two countries (i.e. without the assumption of a common unobserved heterogeneity distribution). For Sweden, this hardly changed the results at all, while for Norway, the baseline exhibited slightly less negative duration dependence (with positive duration dependence during the first two months).
50 per cent in the months just prior to benefit exhaustion. This corresponds almost exactly to the exhaustion effect reported by Røed and Zhang (2003) regarding the previous 80-week benefit limit in Norway that applied until 1997. Hence, it seems that the limited benefit period in Sweden has a substantial effect on search behaviour, and that this was the case even when benefit entitlements could be renewed through participation in labour market programs. Previous Swedish studies (Carling et al, 1996; Thoursie, 1998) have failed to establish this effect in a convincing way, due to large standard errors in the estimated spell duration baselines (because of relatively small risk sets with long durations). It is interesting to note, however, that the point estimates of the exhaustion effect reported by both Carling et al (1996) and Thoursie (1998) are not far from our own (very precise) estimates. The substantial and significant rise in the job hazard during the months just prior to benefit exhaustion establishes beyond doubt that even the very ‘soft’ benefit duration limit operating in Sweden, affect search behaviour in the intended way. The prospect of being enrolled into a labour market program does apparently not have the perverse effect often identified in search models (see e.g. Carling et al, 1996) of raising reservation wages and lowering search effort. On the contrary, program enrolment seems to be considered a ‘stick’ rather than a ‘carrot’ by many unemployed workers. One obvious explanation is that program participation reduces the quantity of leisure. Our results at this point are in accordance with experimental evidence reported by Black et al (2002)

---

4 We also estimated the model for Sweden with only positively identified job transitions recorded as true transitions (i.e. we censored observations ending with ‘lost contact’). This did not change the results to any large extent, although the estimated rise in the hazard rate in the months just prior to exhaustion was reduced to around 45 per cent. Note that the large confidence intervals after month 14 in the Swedish baseline simply reflects that most of the remaining persons who do not obtain a job transit to labour market programs at this stage of the spell, hence very few are left in the population at risk.
indicating that the ‘threat’ of (mandatory) employment and training services has a much stronger effect on employment transitions than actual participation.

The strong exhaustion effect may not only reflect the prospect of having to participate in labour market programs. An additional explanation is that persons with imminently expiring benefits are given high priority in the allocation of labour market services. Prior to benefit exhaustion the job seekers are typically summoned to consultations at the Public Employment Service in order to discuss employment opportunities, job search efforts, alternative possibilities of income support etc. Dolton and O’Neill (1996) and Gorter and Kalb (1996) report evidence indicating that just by giving attention to unemployed workers (e.g. by summoning them to an interview) employment offices stimulate them to find a job more quickly.
The effects associated with a marginal change in the replacement ratio are also stronger in Norway than in Sweden. The Norwegian elasticity estimate around –1.0 indicates slightly stronger responses than previous results reported for Norway by Røed and Zhang (2003), while the Swedish elasticity estimate of –0.5 indicates weaker responses than previously reported for Sweden by Carling et al (2001). A weaker Swedish disincentive effect could result from a more active Public Employment Service, leaving less room for individuals to optimise freely with respect to their own reservation wage and search intensity. It should be noted in passing that the two hypothetical replacement ratios, which by construction have no causal effect on the hazard rates, are attributed coefficients that are larger (in absolute terms) than the true causal effects (see the bottom of Table 2). This suggests that the failure to control for spurious correlation between replacement ratios and transition rates is likely to produce misleading results.

5 Conclusion

The institutional set-up of the unemployment insurance system may have a substantial influence on the escape rate from unemployment. In the present paper, we have combined register data from Norway and Sweden to identify causal effects associated with the replacement ratio and the maximum benefit duration. We find that structural differences between the two countries are clearly mirrored in the transition rate pattern from unemployment into jobs. While the relatively long maximum benefit duration in Norway (3 years) produces a pattern of monotonously declining employment prospects over spell

\footnote{In our case, it turns out that unobserved heterogeneity tends to bias the elasticities towards zero, particularly for Norway. Without the controls for spurious correlation (i.e. with separate estimations for each country), we obtain elasticity estimates of \(-0.49\) (0.09) in Norway and \(-0.43\) (0.02) in Sweden (standard errors in parentheses).}
duration, the short maximum benefit eligibility period in Sweden (60 weeks) delivers a relatively stable hazard rate, with a sharp increase in the months just prior to benefit exhaustion. This may be viewed as a surprising result, given that benefit entitlements in Sweden could easily be renewed through participation in labour market programs during the observation period. A corollary seems to be that the ‘threat’ of compulsory program participation (or otherwise losing UI benefits) has a strongly encouraging effect on job search efforts. This is a type of ‘program-effect’ that has received little attention in the otherwise extensive literature on program evaluation in Sweden (see e.g. the recent survey by Calmfors et al, 2002).

The difference in the two countries’ replacement ratio profiles is clearly mirrored in the pattern of job transition rates. The ratio of transition rates for low- to high income workers is much smaller in Sweden than in Norway, reflecting that the ratio of replacement ratios for low- to high income workers is much larger in Sweden than in Norway. When we use the between-countries variation in replacement ratios to identify causal effects, we also find that the effects associated with marginal changes in economic incentives are larger in Norway (with a replacement ratio elasticity around –1.0) than in Sweden, (with an elasticity around –0.5). This may suggest that the ‘active’ labour market policies pursued in Sweden leave less scope for individuals to optimise freely with respect to search intensity and reservation wages.

The potentially most important lesson arising from the empirical results presented in this paper is that it indeed seems to be possible to counteract the disincentive effects associated with a generous UI benefit system through a combination of ‘soft’ benefit duration limits and active labour market policies.
References


