

IZA DP No. 632

Duration Dependence in the Exit Rate out of Unemployment in Belgium: Is It True or Spurious?

Bart Cockx
Muriel Dejemeppe

November 2002

Duration Dependence in the Exit Rate out of Unemployment in Belgium: Is It True or Spurious?

Bart Cockx

IRES, Université Catholique de Louvain and IZA Bonn

Muriel Dejemeppe

IRES, Université Catholique de Louvain

Discussion Paper No. 632
November 2002

IZA

P.O. Box 7240
D-53072 Bonn
Germany

Tel.: +49-228-3894-0
Fax: +49-228-3894-210
Email: iza@iza.org

This Discussion Paper is issued within the framework of IZA's research area *Internationalization of Labor Markets*. Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent, nonprofit limited liability company (Gesellschaft mit beschränkter Haftung) supported by the Deutsche Post AG. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public. The current research program deals with (1) mobility and flexibility of labor, (2) internationalization of labor markets, (3) welfare state and labor market, (4) labor markets in transition countries, (5) the future of labor, (6) evaluation of labor market policies and projects and (7) general labor economics.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website (www.iza.org) or directly from the author.

ABSTRACT

Duration Dependence in the Exit Rate out of Unemployment in Belgium: Is It True or Spurious?

On the basis of aggregate data for the early nineties, we analyse the determinants of unemployment duration for laid-off male workers in Wallonia (Belgium). Our results demonstrate that if ranking in recruitment occurs, the standard Mixed Proportional Hazard specification can be too restrictive, leading to an overstatement of the extent of true negative duration dependence. We conclude that negative duration dependence is largely spurious. We also decompose the time variation of the hazard in (unobserved) compositional and direct cyclical and seasonal effects. We find counter-cyclical variation in the quality of young workers, but none for the prime aged.

JEL Classification: C41, J64

Keywords: unemployment duration, ranking, heterogeneity, business cycle

Corresponding author:

Bart Cockx
Institut de Recherches Economiques et Sociales (IRES)
Université Catholique de Louvain
Place Montesquieu 3
1348 Louvain-la-Neuve
Belgium
Tel.: +32 (0)10 473439
Fax: +32 (0)10 473945
Email: cockx@ires.ucl.ac.be

1 Introduction

Finding an explanation for the rise and persistence of European unemployment has been one of the main research programmes of labour economists in the last decade. This study contributes to this literature by decomposing unemployment duration in Wallonia, the French speaking part of Belgium¹, into its determinants. The case of Belgium is particularly interesting, since the share of long-term unemployment in this country is one of the highest in Europe. During the 1983-1996 period this share peaked at a level of 68%² and has only been slightly below 60% in the recession of the early nineties.

The primary objective of this paper is to investigate whether duration dependence in the exit rate from unemployment in Wallonia is true or spurious. Although these questions are old (see Salant 1977, Nickell 1979, Lancaster 1979), disentangling true from spurious duration dependence remains highly relevant from a policy point of view. For continental European countries researchers generally do not find evidence of marked negative duration dependence once observed and unobserved heterogeneity is controlled for³ (for a recent review, see Machin and Manning 1999). Belgian studies (see Spinnewyn 1982, Plasman 1993 and Mahy 1994) report that negative duration dependence is completely spurious. The hazard rate increases significantly once unobserved heterogeneity is taken into account. However, in view of the strong parametric assumptions regarding both the baseline hazard (Weibull) and the mixing distribution (Gamma), these results are possibly biased.

A second aim of this paper is to study the calendar time behaviour of the exit rate out of unemployment in Wallonia. The aggregate outflow rate is generally found to be procyclical. This dynamic fluctuation can be driven by variations in the quality of entrants into unemployment rather than by the effect of the business cycle on the exit rate of all currently unemployed workers. In this paper, we disentangle the direct effect of the business cycle and seasons on the hazard from its compositional effect.⁴ However, since the observation period doesn't even span a complete cycle, this decomposition will not be the main focus of this paper. This is the main goal of Dejemeppe and Saks (2002) analysing annually⁵ grouped data on the exit rate in Belgium covering a period of 21 years. These authors report that compositional changes affect the cyclical fluctuations in the aggregate hazard only marginally, but drive an important part of its long run dynamics. Given the recent interest of European researchers in this question, only a few

¹Belgium is a federal state delegating a fair amount of autonomy to three regions: Wallonia, Flanders and Brussels.

²This figure gives the share of workers who have been unemployed more than one year. Workers are defined to be unemployed if they are entitled to unemployment benefits and actively looking for a job.

³In contrast, researchers generally report that in the United Kingdom the conditional hazard displays strong negative duration dependence.

⁴Compositional effects cannot be distinguished from competition effects induced by variations in the size of the entry cohort (Abbring *et al.* 2002). In order to facilitate discussion, we ignore competition effects in the sequel. The reader can always interpret the results in terms of competition effects, however.

⁵Given the annual periodicity of these data, they are less suited to analyse duration dependence. The two studies are therefore complementary.

researchers have studied the dynamics of unemployment duration in Europe.⁶ Abbring *et al.* (2002) and Rosholm (2001), respectively for France and Denmark, conclude that compositional effects are of minor importance. On the other hand, Kalwij (2001) finds that changes in the quality of entrants account for a significant part of the cyclicity in the aggregate outflow rate for the United Kingdom.

The analysis of this study is based on aggregate data. We observe the quarterly exit rates of male workers flowing into unemployment between June 1989 and December 1993. Using this type of aggregate data, Abbring, Van den Berg and Van Ours (2002) developed an estimation method in which both the distribution of duration and of unobserved individual characteristics are specified non-parametrically. They prove that the model is identified if the underlying data are generated by a discrete time process and if the hazard is of the mixed proportional hazard (MPH) form.

Abbring *et al.* (2002) assume that the data are generated by a discrete time process. It is well known (see e.g. Flinn and Heckman 1982, pp.53-56) that with such an assumption the parameters are not invariant to the time unit. Consequently, consistency requires that the timing of the underlying stochastic duration process coincide with the actual grouping of the data.⁷ In order to avoid this problem, we specify the discrete-time process as a continuous-time model (see e.g. Van den Berg and Van der Klaauw, 1998). As a consequence, we cannot apply the identification results and the estimation method proposed by Abbring *et al.* (2002).⁸

Our analysis exploits Ridder's (1990) identification results for grouped data within a continuous-time MPH model. Nevertheless, we relax the MPH assumption in two minor ways. First, we allow the baseline hazard to vary non-proportionally between two sub-periods. The reason for doing so is that if true negative duration dependence is caused by ranking, a recruitment rule that consists in hiring the candidate with the shortest unemployment duration, the baseline hazard is predicted to vary non-proportionally over the business cycle: It will decline more rapidly the more depressed is the labour market (see Blanchard and Diamond 1994). Our results demonstrate that a neglect of this time-duration interaction term may lead to an over-estimation of true negative duration dependence. Identification is ensured by maintaining the MPH assumption within the two sub-periods.

Second, the MPH specification assumes that the business cycle at the time of exit influences the hazard of all cohorts, i.e. of all duration classes at this instant, with the same factor of proportionality. We generalise by allowing for random deviations from this specification, i.e. by introducing cohort-specific random business cycle effects at the time of exit. We show that for our data the correlation structure of the residuals is compatible with this type of specification error.

⁶This question has puzzled American researchers since the mid 1980s already (see e.g. Darby *et al.* 1985, Sider 1985, Dynarski and Sheffrin 1990, Baker 1992, Imbens and Lynch 1992).

⁷The authors are aware of this problem and test the validity of this assumption (see e.g. Van den Berg and Van Ours 1994, p.437).

⁸Another reason to prefer the estimation method proposed in this paper is that its computational complexity does not increase with the number of duration intervals, in contrast to the one used by Abbring *et al.* (2002).

We estimate the duration model by Minimum Chi-Squares (Berkson 1944, Amemiya 1981; for an application on transition data, see Cockx 1997). We prefer this method to Maximum Likelihood, since it can more easily accommodate random specification errors of the above-mentioned type or measurement errors (see Amemiya and Nold 1975, Cockx and Ridder 2001) and is therefore more robust to misspecification.⁹

The baseline hazard is specified as piecewise-constant. Related to our objective to disentangle compositional from direct effects of calendar time on the outflow rate, we allow the scale of the mixing distribution to vary over seasons and the business cycle. In the benchmark model the mixing distribution is Gamma. The findings of this model are contrasted to those obtained with a non-parametric specification of the heterogeneity distribution, as approximated by one with a discrete number of points of support.

The plan of the paper is as follows. The next section describes the data, introduces the main notation and specifies the model. The third section explains the estimation method. The estimation results are discussed in the fourth section. The concluding part summarises our findings.

2 Data, notation and specification

Our analysis exploits quarterly census data relative to the unemployed male population in Wallonia (Belgium) and stratified by the quarter of inflow, the unemployment duration and the age group (≤ 28 years old, and 29-44 years old).

Workers are defined to be unemployed if they are officially registered as full-time unemployed and if a sufficiently long employment record entitles them to unemployment benefits. The inflow in a given quarter is equal to the number of laid-off workers who are still unemployed at the end of the quarter in which they enter.¹⁰ The observation period consists of 19 quarterly intervals,¹¹ the first interval starting on the 1th of June 1989, the last one ending on the 28th of February 1994. So there are 18 cohorts of unemployed¹² (a cohort being defined by its quarter of inflow), stratified by age group. For each cohort, the flows out of unemployment are counted on a quarterly basis, from the time of inflow until the end of February 1994, the date at which all spells are right censored. For each age group, the data are therefore grouped into $N = \sum_{n=1}^{18} n = 171$ homogeneous cells.

⁹The Minimum Chi-Square Method explicitly allows the empirical exit probability to deviate randomly from its theoretical counterpart in the super-population. As the size of the population at risk (of leaving unemployment) varies over calendar time and duration, this approximation error term is heteroskedastic. This is neglected by Abbring *et al.* (2002). Their approach can be justified in two ways. First, their inference is valid for the particular population they observe rather than for the hypothetical super-population. Second, if there are many unemployed in each duration interval, the approximation errors will be small. They are likely to be larger in our study since we consider a smaller population and up to 18 quarters of unemployment, whereas Abbring *et al.* (2001a) deal with only up to four duration quarters.

¹⁰An exit should also last at least three months to be recorded as such. This restriction aims at restricting the number of fictive exits (for administrative reasons for example).

¹¹Four quarters are defined: (Jun,Jul,Aug); (Sep,Oct,Nov); (Dec,Jan,Feb); (Mar,Apr,May).

¹²There are 18 cohorts instead of 19 as there is no exit defined for entrants in the 19th quarter.

The data do not allow us to distinguish between unemployment spells ending in employment, participation in labour market programmes or withdrawal from the labour force. This is not likely to bias our results for two reasons. First, employment is the most frequent state of exit for the male unemployed workers: About 77% of them become employed,¹³ against 47% for women (see FOREM 1995). Second, the period of entitlement to benefits is in principle indefinite in Belgium.

In the sequel, the variables l and t denote respectively the calendar time at the moment of inflow and unemployment duration. These variables are measured in quarters. The way the inflow is constructed in our data requires us to model calendar time at the moment of inflow, l , as a discrete time process. We have that $l \in \{0, 1, \dots, 17\}$ and by definition, that $t \equiv 0$ at the moment of inflow l . Although we analyse discrete data, we treat unemployment duration, t , as a continuous time process (as in Van den Berg and Van der Klaauw 2001). So, $t \in R^+$.

In this paper, we depart from the mixed proportional hazard (MPH) specification imposed by Abbring *et al.* (2002). The baseline hazard is allowed to vary non-proportionally with the business cycle so as to control for the ranking hypothesis and therefore avoid under-estimation of the extent of spurious negative duration dependence. In the ranking model, the effects of duration on the baseline hazard are stronger the more depressed is the labour market (see Blanchard and Diamond 1994). As the state of the labour market deteriorates, the long term unemployed are more rapidly excluded from the recruitment process as compared to the situation they face in a tight labour market. However, the identification of unobserved heterogeneity, v , in the MPH model comes from the negative interaction between the duration dependence of the aggregate hazard (i.e. unconditional on v) and calendar time. Heterogeneity always implies that the effects of duration on the aggregate hazard are stronger in a tight labour market (see Van den Berg 2001). Since the time-duration interaction generated by ranking goes in the other direction, the MPH assumption will lead to an under-estimation of unobserved heterogeneity if ranking occurs.¹⁴

For identification purposes, we cannot allow for an interaction between the baseline hazard and all calendar time intervals, but only for two sub-periods.¹⁵ In order to contrast the two periods considered, the first one covers the peak observed between September 1989 and November 1990, and the second one, the slack period starting in December 1990 and continuing up to February 1994.

We denote the hazard of an individual unemployed for t quarters and at risk of leaving unemployment at calendar time $l + t$, conditional on unmeasured characteristics v , by $h(t | l + t, v)$. We impose, without loss of generality, that both the baseline hazard and

¹³In our sample, this fraction must be larger. For, the reported figure includes unemployed workers older than 44 years, who often enter a kind of early-retirement scheme.

¹⁴Abbring *et al.* (1997) argue that the MPH assumption holds if the time-duration interaction in the aggregate hazard is significantly negative. However, a negative sign only implies that sorting dominates ranking. Both processes can operate simultaneously.

¹⁵Kalwij (2001) and Rosholm (2001) allow for a full interaction between the baseline hazard and the business cycle. Since they observe multiple spells for a given individual, identification of unobserved heterogeneity is ensured without imposing restrictions on the time-duration interaction.

the calendar time effects are constant within quarterly intervals. The extended MPH specification then takes the following form:

$$\forall t \in [k-1, k) : h(t|l+t, v) = \exp[\varphi_1(k) + \varphi_2(l+k) + \varphi_3(k, l+k)]v \equiv h_{kl}v \quad (1)$$

where $v > 0$, $k \in \{1, \dots, 18\}$. The functions $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$ represent respectively the duration and calendar time dependence of the conditional hazard. The function $\varphi_3(\cdot, \cdot)$ captures the time-duration interaction.

We parametrise $\varphi_1(k)$ as a piecewise-constant function:

$$\varphi_1(k) = c + \sum_{j=2}^{18} (\gamma_j - c) \delta_{jk} \quad (2)$$

where δ_{jk} is the Kronecker delta and where we impose that $\forall j \geq 12 : \gamma_j = \gamma_{12}$. The latter restriction is imposed to avoid erratic behaviour of the parameters induced by too small a number of observations at the tail of the duration distribution.

In the line of Abbring *et al.* (2001), we represent the calendar time dependence of the hazard, $\varphi_2(l+k)$, by a flexible parametric specification, where we decompose $\varphi_2(l+k)$ into a cyclical part, $\varphi_{2c}(l+k)$, and a seasonal part, $\varphi_{2s}(l+k)$ such that $\varphi_2(l+k) = \varphi_{2c}(l+k) + \varphi_{2s}(l+k)$ and

$$\begin{aligned} \varphi_{2c}(l+k) &= \sum_{i=1}^5 \beta_{ci} p_i(l+k) \\ \varphi_{2s}(l+k) &= \sum_{a=1}^4 \beta_{sa} [s_a(l+k) - s_a(1)] \end{aligned} \quad (3)$$

in which we normalise $\beta_{s1} = 0$. The effect of calendar time at the outflow on the hazard is represented by a flexible fifth degree Chebyshev polynomial in calendar time, $p_i(l+k)$, capturing business cycle effects, and by dummy variables, $s_a(l+k)$, capturing seasonal effects.

The time-duration interaction is specified as follows:

$$\varphi_3(k, l+k) = \phi I_{\{1,2,3,4,5\}}(l+k) \sum_{j=2}^{18} |\gamma_j - c| \delta_{jk} \quad (4)$$

The interaction parameter, ϕ , is only defined for the boom period (i.e. the first five time intervals) and is assumed to affect the logarithm of the baseline hazard in absolute value multiplicatively. We take the absolute value to avoid that the sign of the duration dependence affects the sign of the interaction term. Ranking implies that $\phi > 0$.

We tried out alternative specifications of the interaction effect, $\phi |\gamma_j - c|$,¹⁶ $\phi \log(k)$ and $\phi_1 \log(k) + \phi_2 \log^2(k)$. However, these yield much less precise parameter estimates than our chosen specification without improving the goodness-of-fit.¹⁷

¹⁶This term actually represents the *cumulative* interaction effect up to duration j (see Van den Berg 2001, p.3417).

¹⁷These estimation results can be obtained from the authors on request.

If we assume that the underlying duration process is continuous and if calendar time effects are constants within each quarter, then the mixture survivor function at the end of quarter k is (Lancaster, 1990, p.9 and p.59):¹⁸

$$S^m(k|l) = \int_0^{\infty} \exp \left[-v \sum_{j=1}^k h_{jl} \right] dF_l(v) \quad (5)$$

where $F_l(\cdot)$ is the mixing distribution at the calendar time of inflow l . As in Abbring *et al.* (2002), we assume that the mixing distribution varies with l according to a scale parameter, $\exp[\varphi_4(l)]$, such that

$$F_l[\exp[-\varphi_4(l)]v] = F_0(v) \equiv F(v) \quad (6)$$

where $\varphi_4(0) = 0$ by normalisation. This assumption ensures a MPH specification within the two above-mentioned sub-periods, as required for identification (see below).

We will contrast two specifications for the mixing distribution, $F_l(v)$. In principle, a non-parametric specification would be the best choice. However, recent Monte Carlo analysis by Baker and Melino (2000) reveals that a non-parametric specification of both the distribution of duration and of unobserved heterogeneity can result in large and systematic bias, if estimated by maximum likelihood. Baker and Melino show that the bias can be reduced by modifying the objective function such that specifications with many points of support are penalised. We will follow a similar procedure below, based on the Schwarz (1978) information criterion.

The findings of Baker and Melino suggest that a parametric specification of the mixing distribution could improve upon a non-parametric one. Recently, Abbring and Van den Berg (2001) justify the choice of a Gamma distribution. They show that under mild regularity conditions an unobserved heterogeneity distribution with continuous support converges to a Gamma distribution for $t \rightarrow \infty$. In the benchmark model the mixing distribution is therefore assumed to be Gamma. We then compare our findings with those found with a non-parametric specification of the mixing distribution, as approximated by one with a discrete number of mass points.

Assumption (6) implies that the moment generating function of the Gamma variate \mathbf{v} with mean, $\mu_l = \frac{\lambda_l}{\delta_l}$, and variance, $\sigma_l^2 = \frac{\lambda_l}{\delta_l^2}$, must satisfy the following restriction:

$$\left(1 - \frac{s \exp[-\varphi_4(l)]}{\delta_l} \right)^{-\lambda_l} = \left(1 - \frac{s}{\delta_0} \right)^{-\delta_0} \quad (7)$$

in which the mean is normalised to one at $l = 0$: $\lambda_0 = \delta_0$. This restriction implies that $\lambda_l = \delta_0$, and $\delta_l = \exp[-\varphi_4(l)]\delta_0$. As such, the mean and variance of the mixing distribution are required to vary proportionally over calendar time at inflow.

For a discrete mixing distribution with S points of support it is easy to verify that assumption (6) implies the mass points, v_{sl} , and the corresponding probabilities, π_{sl} ,

¹⁸Note that, if the underlying duration process were discrete, then the specification of the survivor function would differ (Abbring *et al.* 2002).

must satisfy the following restrictions:

$$\begin{aligned} v_{sl} &= v_s \exp[\varphi_4(l)] \\ \pi_{sl} &= \pi_s \end{aligned} \tag{8}$$

for $\forall s \in \{1, 2, \dots, S\}$ and where $v_s > 0$ and $0 \leq \pi_s \leq 1$. The mean is normalised to one for $l = 0$ by imposing $v_1 = \frac{1 - \sum_{s=2}^S \pi_s v_s}{\pi_1}$.

Finally, we complete our specification by parameterising $\varphi_4(l)$. We decompose it into a cyclical ($\varphi_{4c}(l)$) and a seasonal component ($\varphi_{4s}(l)$) such that $\varphi_4(l) = \varphi_{4c}(l) + \varphi_{4s}(l)$. In contrast to the effect of calendar time on the hazard at the moment of exit, we choose to characterise the effect of calendar time at entry by a business cycle indicator and by the season at inflow:

$$\begin{aligned} \varphi_{4c}(l) &= \eta_c [b(l) - b(0)] \\ \varphi_{4s}(l) &= \sum_{a=1}^4 \eta_{sa} [s_a(l) - s_a(0)] \end{aligned} \tag{9}$$

where $s_a(l)$ are the seasonal dummies at the time of inflow and where we normalise $\eta_{s1} = 0$. Note that the specification ensures that $\varphi_4(0) = 0$.

The business cycle indicator, $b(l)$, is the de-seasonalised logarithm of the number of quarterly flows into unemployment. We choose this regressor rather than a polynomial in calendar time at entry, because it allows us to identify a trend in the unobserved composition effect. In fact, an exponential trend in $\varphi_{4c}(l)$ cannot be non-parametrically identified from the exponential trend in the calendar time effect at outflow, β_{c1} , (see Abbring *et al.* 2002). Identification follows from the assumption that the variation of $b(l)$ around or along this trend evenly affects the scale of the mixing distribution.

Note that by assumption (6), $b(l)$ and $s_a(l)$ can be viewed as additional regressors affecting the baseline hazard proportionally. The identification of the mixture distribution comes therefore not only from the interaction effects of duration with the calendar time at exit within the upturn and the downturn (cf. *supra*), but also from the interaction with these additional regressors. Consequently, one could question why we don't allow for a full interaction between calendar time at exit and duration. As such, there could be ranking within sub-periods, identifying heterogeneity by the interaction with the remaining regressors. The reason is that these time-varying regressors are not independent. For any given combination of calendar time and duration, we observe, by definition, only one cohort and accordingly, no variation in $\varphi_4(l)$ to identify heterogeneity. This is what differentiates our analysis from the standard MPH with time-constant regressors.

Grouped duration data require stronger identifying assumptions than continuous duration data. Within a MPH model, a sufficient condition for non-parametric identification is the regression function takes on every value in R instead of only two distinct values (see Ridder, 1990). With our specification of $(\varphi_2(l+k) + \varphi_4(l))$ this requirement

is only approximated. We believe that this is not troublesome in practice, since the specification of the mixing distribution is not required to vary in an unrestricted way over the unknown continuous support. It is restricted, either to one with a relatively small number of mass points or to a member of a parametric family determined by a small number of parameters (cf. supra).

3 The mixture models

We estimate our model by Minimum Chi-Square (see Berkson 1944, Amemiya 1981, Cockx 1997). In a nutshell, this method consists in regressing, for each cohort, the observed exit probability to its theoretical counterpart. The theoretical probability at the k^{th} duration interval for a worker who entered unemployment at the l^{th} quarter is:

$$P_{kl} \equiv \Pr(k-1 \leq T < k | T \geq k-1, l+k) \quad (10)$$

Using (5), we then link this probability to the conditional hazard defined in (1):

$$P_{kl} = \frac{S^m(k-1|l) - S^m(k|l)}{S^m(k-1|l)} = 1 - \frac{\int_0^\infty \exp\left[-v \sum_{j=1}^k h_{jl}\right] dF_l(v)}{\int_0^\infty \exp\left[-v \sum_{j=1}^{k-1} h_{jl}\right] dF_l(v)} \quad (11)$$

We now relate the theoretical probability to the empirical one. Let u_{kl} be the number of individuals who entered into unemployment at l , are still in that state at the start of the k^{th} duration interval and are therefore at risk of leaving unemployment within the $(l+k)^{th}$ quarter. Let f_{kl} be the number of these individuals who leave unemployment within the k^{th} duration interval. We can then estimate the probability of leaving unemployment, P_{kl} , by the Kaplan-Meier estimator of the hazard adapted to grouped data:

$$\hat{P}_{kl} = \frac{f_{kl}}{u_{kl}} \quad (12)$$

In the benchmark model we assume a Gamma mixing distribution. Combining (7), (11) and (12) then yields the following non-linear heteroskedastic regression model (see Cockx 1997):

$$1 - \hat{P}_{kl} = \left[1 - \frac{\exp[-\delta_0^* + \varphi_4(l)] h_{kl}}{1 + \exp[-\delta_0^* + \varphi_4(l)] \sum_{j=1}^k h_{jl}} \right]^{\exp[\delta_0^*]} + \omega_{kl} + \varepsilon_{kl} \quad (13)$$

where $\delta_0^* = \ln(\delta_0) \equiv \ln(\sigma_0^{-2})$ and $\omega_{kl} = \hat{P}_{kl} - P_{kl} - \varepsilon_{kl}$ is the approximation error. The errors ε_{kl} are specification errors. We assume that they have a distribution such that $E(\varepsilon_{kl}) = 0$ and $E(\varepsilon_{kl}^2) = s_\varepsilon^2$. We have that $E(\omega_{kl}) = 0$ and its variance can be consistently estimated by (see Cockx 1997):

$$\hat{s}_{kl}^2 = \left[\left(\hat{P}_{kl} \right) \left(1 - \hat{P}_{kl} \right) \right] \frac{1}{u_{kl}} \quad (14)$$

The correlation structure of the specification errors is informative about the type of misspecification. We find that the empirical correlation between the residuals of adjacent duration classes for any fixed calendar time at the outflow, i.e. the correlation between ε_{kl} and ε_{k-1l+1} , denoted by ρ_{l+k} , is significantly positive according to the LM-test statistic for both the men aged 29-44 and those aged ≤ 28 : respectively, 0.65 and 0.77 (see Appendix 1 for their measure and test). These results suggest that these residuals reflect cohort-specific random business cycle effects at the time of exit. On average these random deviations from the MPH specification cancel out. However, to the extent that cohorts with a different elapsed duration are imperfect substitutes, a variation in the hazard of one particular cohort will spill over to the other cohorts. If we assume that the degree of substitutability decreases with the difference of the elapsed duration between the affected and non-affected cohorts, then this substitution induces, for any fixed exit time, the residuals to be positively autocorrelated with duration.

Note that ε_{kl} could also have been interpreted as a measurement error in the line of Abbring *et al.* (2002). Measurement errors always imply that the correlation between the residuals of adjacent duration classes for any fixed calendar time at the inflow, i.e. the correlation between ε_{kl} and ε_{k-1l} , denoted by ρ_l , is equal to $-\frac{1}{2}$. For our data, the empirical correlation between ε_{kl} and ε_{k-1l} is, however, not significantly different from zero according to the LM-test statistic, whatever the age group.

We estimate the statistical model (13) by Generalized Non-Linear Least Squares. The estimation procedure consists of two-steps (see Amemiya and Nold 1975, Cockx and Ridder 2001) and is detailed in Appendix 2.

Because we have grouped data, we can use a χ^2 -goodness-of-fit test to evaluate the model specification. The weighted sum of squared residuals (*WSSR*) is distributed $\chi_{N-N_\theta}^2$ if the model is correctly specified, where N denotes the number of cells and N_θ the number of parameters. Moreover, we can test the acceptability of parameter restrictions. The difference between the weighted sum of squared residuals of the constrained and the unconstrained models ($WRSS_0 - WRSS$) is distributed $\chi_{(N_\theta - N_{\theta_0})}^2$.¹⁹

We contrast the findings of the benchmark model to those in which the mixing distribution is non-parametrically distributed. Heckman and Singer (1984) propose a procedure to estimate such a model by maximum likelihood, assuming that the unobserved random effects are drawn from a discrete distribution with unknown support and an unknown number of mass points. Baker and Melino (2000) supply an heuristic interpretation and adjust this procedure to reduce the bias of the estimator in finite samples. They propose to maximise the Hannan-Quinn or the Schwartz Information (SI) criterion rather than the log-likelihood function (see also Leroux 1992). They report (footnote 14, p.376) that results are virtually identical for these two criteria. We choose the SI criterion (see Chow 1985, pp.300-305, for a justification).

¹⁹For the purpose of this test statistic, the constrained model is to be estimated with the estimated variance matrix of the specification errors in the unconstrained model.

To make the above mentioned procedure applicable, we need to modify the objective function, i.e. we minimise the *WSSR* rather than maximise a loglikelihood function. How this modifies the rule is discussed in Appendix 3. In the SI criterion, the *WSSR* replaces minus twice the loglikelihood value (see Amemiya 1981). According to this criterion, the preferred model is therefore the one that minimises the following statistic:

$$WSSR + N_{\theta} \ln(N) \quad (15)$$

We use this criterion not only to determine the number of mass points, but also to judge the performance of the Gamma to the non-parametric specification.

By combining (8), (11) and (12) we obtain the non-linear heteroskedastic regression model for a discrete mixing distribution with S points of support:

$$1 - \hat{P}_{kl} = \frac{\sum_{s=1}^S \pi_s \exp \left[-v_s \exp [\varphi_4(l)] \sum_{j=1}^k h_{jl} \right]}{\sum_{s=1}^S \pi_s \exp \left[-v_s \exp [\varphi_4(l)] \sum_{j=1}^{k-1} h_{jl} \right]} + \omega_{kl} + \varepsilon_{kl} \quad (16)$$

4 Estimation results

4.1 Specification tests

In Tables 1 and 2 we report the estimation results, respectively for men aged 29-44 (Table 1) and for those aged ≤ 28 (Table 2). The estimates of the Gamma and the non-parametric mixture model²⁰ are reported, respectively, in the first and last column. For the young men, we also report the Gamma mixture with a constant baseline hazard and a MPH assumption (i.e. $\phi = 0$) (third column of Table 2). According to the χ^2 -test, the latter restrictions are not rejected at a significance level of 5% (P-value= 36%). For the young group, we therefore estimate the non-parametric mixture model on which we impose a constant baseline hazard and $\phi = 0$ (last column of Table 2).

INSERT TABLES 1 AND 2 APPROXIMATELY HERE

All mixture models fit the data reasonably well. According to the goodness-of-fit statistic, no model, except the Gamma mixture model for young unemployed workers, can be rejected against the saturated model at a significance level of 5%. In the procedure that selects the number of mass points of the non-parametric mixing distribution, we can decrease the *WSSR* marginally by going up from two to three points. The probability assigned to the third point is very small for the prime aged men while it is significantly positive for the young men. However, on the basis of the SI criterion, the model with two points of support is preferred for both age groups.

²⁰The reported standard errors are the usual ones, conditional on the number of points of support. These do not have any rigorous justification, however, since the number of points of support is also a parameter to be estimated and can therefore not be conditioned upon (Heckman and Singer 1984, p.300).

As to the estimation of mixture models, our analysis is in line with the recent Monte Carlo evidence of Baker and Melino (2000). For the young workers, a specification with constant baseline hazard cannot be rejected. In that case, Baker and Melino report that the non-parametric mixture performs well. According to the SI criterion, we indeed reject the parametric Gamma mixture against the non-parametric one. On the other hand, an exponential baseline hazard is rejected for the older workers. In that case, Baker and Melino report that parameter estimators are biased in finite samples. Even if their modified estimation procedure based on the SI criterion leads to a dramatic improvement, it does not guarantee a better performance as compared to a parametric mixing distribution. In this application, the Gamma mixture indeed performs better than the non-parametric one for the older workers. In the following discussion, we retain the preferred model, i.e. the discrete mixture for the young men and the Gamma for the prime aged men.

Finally, note that the specification error accounts for only 1% of the variation in the dependent variable for the adult population while it is somewhat larger for the young unemployed workers (3%). Cohort-specific calendar time effects are thus unimportant. The approximation error amounts on average to only 1% of the variation in \hat{P}_{kl} for the first four duration quarters while it accounts for more than 7% for long durations (>4 quarters).

4.2 True versus spurious duration dependence

The results for the estimated variance of the mixing distribution reveal an important disparity between the characteristics of those flowing into unemployment. Since the variance is allowed to fluctuate over time, we report its average estimated value over our observation period. For prime aged men, the estimated variance of the Gamma distribution is significantly positive (0.89). The dispersion of the two mass points is lower for the young unemployed workers, although significantly different from zero (0.26)²¹.

INSERT FIGURE 1 APPROXIMATELY HERE

Figure 1 shows, for men aged 29-44, the duration pattern of the hazard resulting from three different model specifications: a proportional hazard model neglecting unobserved heterogeneity,²² a MPH model and the extended MPH model (13). In the latter model, the interaction effect ϕ is significantly positive. The duration dependence is positive in the upturn and negative in the downturn. This provides evidence of ranking during the downturn. Consequently the model without interaction underestimates, as expected, the extent of unobserved heterogeneity and overestimates therefore the extent of ‘true’ negative duration dependence.

²¹Note however that the variances of the two age groups are not readily comparable as they come from different distributions. The estimated variance of the Gamma distribution is 0.55 for the young men.

²²In order to take inflow effects into account, we assume that the conditional hazard is multiplicative in the business cycle indicator $b(l)$ and the seasonal dummies at the moment of inflow.

The negative duration dependence in the aggregate hazard in Wallonia is thus largely spurious for men aged 29-44. In a downturn, there is a 7% decrease in the conditional hazard between the first and the second quarter of unemployment, and then a slight increase up to the fifth quarter where the hazard nearly reaches its initial level. Given the level of parameters significance, we can argue that individual exit probabilities are, if anything, slightly decreasing over the first 1.5 years of unemployment and exhibit significant negative duration dependence afterwards. After two and a half years of unemployment, the conditional hazard drops by 20%.

For young men, the observed negative duration dependence of the aggregate hazard is completely spurious. As stated above, the individual exit rate cannot be rejected to be constant over all durations, whatever the state of the business cycle.

How can we explain the differences in the duration pattern between the two age groups? First, in contrast to the adult unemployed, there is no evidence of ranking for the young people. Why is it so? The depreciation of human capital during unemployment may be more important for older workers than for younger ones, because the skills that older workers have acquired on their previous job may require more maintenance than the general skills learnt at school. The reliability of the ranking rule for the prime aged men is reinforced, because their exit rate exhibits true negative duration dependence after 1.5 years that is probably not completely explained by ranking.²³ The higher rate of human capital depreciation can also lead to a more rapid discouragement of older workers in their search for a job.

Differential participation in active labour market programmes (ALMPs) can also drive a different duration dependence between the two age groups. For, participation in ALMPs in Belgium is often targeted to young and long-term unemployed workers. This may mitigate the negative duration dependence in the baseline hazard of the younger workers, since in our data set participants in these programmes are recorded to leave unemployment²⁴ if they do not return within the next quarter.

Another reason is related to temporary exits to employment. In our data set, an exit should last at least three months to be recorded as such. As a consequence, the measure of the elapsed duration into unemployment will not be affected by short exits to employment. Since young workers are over-represented in temporary contracts (see Van Haeperen 2000), they are likely to accumulate short stays into employment during their unemployment spell, thereby maintaining skills and work habits. We might therefore underestimate the extent of true negative duration dependence for young men.

We finish the discussion of our estimation results on duration dependence and heterogeneity by comparing them to findings in recent European studies applying a MPH framework and a flexible functional form for the duration dependence (for a more complete survey, see Machin and Manning 1999).

In view of our results, Belgium belongs to the large pool of European countries where

²³We cannot infer this from our data, because the observation period in the upturn is too short. In the absence of negative duration dependence in the upturn, the negative duration dependence in recession would be completely induced by ranking.

²⁴An exception to this rule is the participation in counseling programmes.

there is both no marked evidence of ‘true’ negative duration dependence of the exit rate out of unemployment and important ‘spurious’ dependence. The United Kingdom remains the exception, with individual unemployment duration distributions exhibiting strong negative duration dependence and not being significantly heterogeneous (see Nickell 1979, Van den Berg and Van Ours 1994, Kalwij 2001). Why we observe this atypical pattern for the UK remains an open question.

With respect to the prime aged men, the shape of the baseline hazard in Wallonia resembles much the one that is found for the Netherlands (Van den Berg and Van Ours 1994 and Kerckhoffs *et al.* 1994) and France (Van den Berg *et al.* 1998). These countries are characterised by non-monotonic true duration dependence, first constant, or even slightly increasing, then decreasing. As to young men, our results are in line with most of the studies that estimate a constant, or weakly negative (after some quarters of inactivity), duration dependence of the exit rate (see for France, D’Addio 1998, Van den Berg and Van Ours 1999 and Van den Berg *et al.* 1998).

We should, however, keep in mind that most of the reviewed studies do not allow the pattern of the duration dependence to vary over the business cycle. This potentially biases their findings regarding duration dependence downwards insofar as calendar time is the identifying regressor. Two exceptions are the studies of Kalwij (2001) and Rosholm (2001), which allow for a full interaction between the baseline hazard and the business cycle. Rosholm estimates that the duration dependence of the hazard becomes less negative, but insignificantly so, when unemployment increases for adult male workers, even after controlling for unobserved heterogeneity. He interprets this as evidence against the ranking model. On the contrary, Kalwij found evidence of ranking for men unemployed aged 18-59. The baseline hazard decreases faster in a downturn than in an upturn, but the difference is small, particularly at short durations.

Finally, note that according to our study, duration dependence, and so ranking, differs between age groups. The common assumption that age affects the baseline hazard proportionally, may accordingly lead to incorrect conclusions.²⁵

4.3 Decomposition of the aggregate hazard over calendar time

For prime aged men, the mixing distribution does not vary significantly over the cycle. According to the χ^2 -test, $\eta_c = 0$ cannot be rejected at a significance level of 5%. But it is rejected for the young men (P-value= 0.5%). Their average quality (and its dispersion) increases significantly over the observation period,²⁶ suggesting a rising share of young entrants with good re-employment prospects as labour market conditions deteriorate in Wallonia.

Comparing the time evolution to a Belgian business cycle indicator (the Kredietbank indicator), the conditional hazard is clearly procyclical for men in Wallonia (Figure 2). The exit rate of the young unemployed was more heavily hit by the recession in 1993

²⁵Rosholm (2001) and Van den Berg *et al.* (1998) specify a distinct baseline hazard by age groups.

²⁶Note that such a statement could not have been made if we had specified the composition effects according to a polynomial function, orthogonal to an exponential trend (cf. the discussion at the end of Section 2).

than the one of the adult unemployed. There is a nearly 40% (resp. 60%) drop in the exit rate in August 1993 compared to its level in November 1989 for the men aged 29-44 (resp. men aged ≤ 28). This is consistent with the lower turnover costs for younger workers.

INSERT FIGURE 2 APPROXIMATELY HERE

With respect to the prime aged men, our findings are in line with the results obtained with French data by Abbring *et al.* (2002) and Van den Berg and Van der Klaauw (2001). They conclude that compositional effects are of minor importance to explain variations in the aggregate outflow rate over the cycle. By contrast, Van den Berg *et al.* (1998) report that the re-employment prospects of young male workers are less affected by the cycle than those of adult workers in France. However, the authors did not account for cyclical composition effects. The strong decline in the outflow rate of young men in recession (see Figure 2) would indeed be mitigated if we did not account for the increase in the average quality at entry into unemployment. Finally, note that we only observe data in the descending phase of a business cycle. We should therefore be cautious in generalising our findings to the complete cycle.

Unlike cyclical effects, seasonal effects induce significant fluctuations of the mean and the variance of the mixing distribution for both age groups. For the oldest men, the average characteristics of new entrants slightly worsen in autumn and spring compared to those observed in the other seasons. The quality of the young entrants is far better in the summer than in all other seasons. We also find significant seasonal variations in the outflow rate that remain unexplained by the seasonal variation in the quality of the inflow. These are about the same for both age groups. The summer and the winter time impedes a lot the exit from unemployment, while the labour market conditions are more favourable in the autumn and in the spring.

5 Conclusion

In this paper, we decomposed the aggregate exit rate out of unemployment in Wallonia into its determinants. We departed from the standard MPH specification by allowing the baseline hazard to vary non-proportionally between the upturn and the downturn. This allowed us to capture ranking in the recruitment behaviour of employers. Related to the objective of disentangling compositional from direct effects of calendar time on the outflow rate, we allowed the scale of the mixing distribution to vary over seasons and the business cycle. We estimated our model by Minimum Chi-Squares on quarterly exit rates of male workers flowing into unemployment between June 1989 and February 1994. We distinguished between two age groups, workers aged less than 29 and between 29 and 44 years old.

We conclude that the negative duration dependence of the aggregate exit rate is largely spurious. For prime aged, but not for young workers, true duration dependence varies over the cycle, a finding consistent with ranking. For young workers, both seasonal

and cyclical variations in the composition of entrants significantly affect their exit rate. While seasonal changes are important, variations in the average quality of prime aged entrants over the cycle is small and insignificant.

On the basis of these results, we conclude that a deterioration of skills and demotivation during unemployment, even if they play a role, cannot be the main channels of unemployment persistence in Wallonia. This is consistent with the findings for a large pool of continental European countries. The pattern of duration dependence in Wallonia resembles much the one reported for the Netherlands and France.

For prime aged men, we significantly improve the fit of our statistical model in departing from the ‘pure’ MPH assumption. Our results demonstrate that, if ranking is important, the MPH assumption imposed on aggregate unemployment duration data may lead to a serious under-estimation of the extent of unobserved heterogeneity. However, the extension of the MPH model proposed in this paper is unsatisfactory. If one allows for a time-duration interaction between two sub-periods, there is no reason for this interaction to be absent within these sub-periods. Accounting for the latter interaction was, however, not possible with our data, since in this case unobserved heterogeneity can no longer be identified. Further research is therefore called for. One strategy would consist in analysing more disaggregated data, including additional continuous regressors. However, identification requires that they enter the specification proportionally and such an assumption can be as imperfect as assuming the absence of time-duration interaction within sub-periods. Another strategy would base the analysis on repeated spells, not requiring proportionality for purposes of identification (as in e.g. Rosholm 2001 and Kalwij 2001).

References

- Abbring, J.H. and G.J. van den Berg (2001), 'The unobserved heterogeneity distribution in duration analysis', Working Paper, Free University Amsterdam, Amsterdam.
- Abbring, J.H., G.J. van den Berg, and J.C. van Ours (1997), 'Business cycles and compositional variation in U.S. unemployment', Discussion Paper n°1702, Centre for Economic Policy Research, London.
- Abbring, J.H., G.J. van den Berg, and J.C. van Ours (2001), 'Business cycles and compositional variation in U.S. unemployment', *Journal of Business and Economic Statistics*, 19 (4), 436-448.
- Abbring, J.H., G.J. van den Berg, and J.C. van Ours (2002), 'The anatomy of unemployment dynamics', *European Economic Review*, 46, forthcoming.
- Amemiya, T. (1981), 'Qualitative response models: A survey', *Journal of Economic Literature*, 19, 1483-1536.
- Amemiya, T. and F. Nold (1975), 'A modified logit model', *Review of Economics and Statistics*, 57, 255-57.
- Baker, M. (1992), 'Unemployment duration: Compositional effects and cyclical variability', *American Economic Review*, 82, 313-321.
- Baker, M. and A. Melino (2000), 'Duration dependence and nonparametric heterogeneity: A Monte Carlo study', *Journal of Econometrics*, 96, 357-393.
- Berkson, J. (1944), 'Application of the logistic function to bio-assay', *Journal of the American Statistical Association*, 39, 357-365.
- Blanchard, O.J. and P.A. Diamond (1994), 'Ranking, unemployment duration, and wages', *Review of Economic Studies*, 61, 417-434.
- Breusch, T.S. and A.R. Pagan (1980), 'The Lagrange multiplier test and its applications to model specification in econometrics', *Review of Economic Studies*, 47, 239-253.
- Chow, G.C. (1985), *Econometrics*, Mc Graw-Hill.
- Cockx, B. (1997), 'Analysis of transition data by the Minimum Chi-Square method: An application to welfare spells in Belgium', *The Review of Economics and Statistics*, 79 (3), 392-405.
- Cockx B. and G. Ridder (2001), 'Social Employment of Welfare Recipients in Belgium: An Evaluation', *The Economic Journal*, 111, 1-31.
- D'Addio, A.C. (1998), 'Unemployment durations of French young people', Discussion Paper n°9821, Institut de Recherches Economiques et Sociales, Université Catholique de Louvain.
- Darby, M.R., J. Haltiwanger, and M. Plant (1985), 'Unemployment rate dynamics and persistent unemployment under rational expectations', *American Economic Review*, 75, 614-637.
- Dejemeppe, M. and Y. Saks (2002), 'A new light into regional unemployment disparities in Belgium: Longitudinal analysis of grouped duration data', Discussion Paper n°2002-19, Institut de Recherches Economiques et Sociales, Université Catholique de Louvain.

Dynarski, M. and S.M. Sheffrin (1990), 'The behavior of unemployment durations over the cycle', *Review of Economics and Statistics*, 72, 350-356.

Flinn, C.J. and J.J. Heckman (1982), 'Models for the analysis of labor force dynamics', in Basmann, R. and G. Rhodes (eds.), *Advances in Econometrics*, 1, 35-95.

FOREM (1995), *Enquête sur les sorties d'inscription des demandeurs d'emploi*, Service Etudes et Statistique, FOREM.

Godfrey, L.G. (1978), 'Testing against general autoregressive and moving average error models when the regressors include lagged dependent variables', *Econometrica*, 46, 1293-1302.

Heckman, J.J. and B. Singer (1984), 'A method for minimizing the impact of distributional assumptions in econometric models for duration data', *Econometrica*, 52, 271-320.

Imbens, G.W. and L.M. Lynch (1992), 'Labor market transitions over the business cycle', *Review of Economics and Statistics*, 72, 350-356.

Kalwij, A. S. (2001), 'Individuals' Unemployment Durations over the Business Cycle', IZA Discussion Paper n°369, Institute for the Study of Labor, Bonn, Germany.

Kerckhoffs, C., C. de Neubourg, and F. Palm (1994), 'The determinants of unemployment and jobsearch in the Netherlands', *De Economist*, 142, 1, 21-42.

Lancaster, T. (1979), 'Econometric methods for the duration of unemployment', *Econometrica*, 47, 939-956.

Lancaster, T. (1990), *The econometric analysis of transition data*, Cambridge University Press, Cambridge.

Leroux, B.G. (1992), 'Consistent estimation of a mixing distribution', *The Annals of Statistics*, 20, 1350-1360.

Machin, S. and A. Manning (1999), 'The causes and consequences of long-term unemployment in Europe', *The Handbook of Labor Economics*, Volume 3C, 3085-3319.

Mahy, M. (1994), 'Politiques et recherche d'emploi: évaluation microéconométrique', *Reflets et Perspectives de la Vie Economique*, 33 (1/2), 87-102.

Nickell, S. (1979), 'Estimating the probability of leaving unemployment', *Econometrica*, 47, 1249-1266.

Plasman, R. (1993), Efficacité des Programmes Publics d'Emploi, Colloque de l'A.E.A. sur l'Evaluation des Politiques d'Emploi, Paris.

Ridder, G. (1990), 'The non-parametric identification of generalized accelerated failure-time models', *The Review of Economic Studies*, 57, 167-181.

Rosholm, M. (2001), 'Cyclical variations in unemployment duration', *Journal of Population Economics*, 14, 173-191.

Salant, S.W. (1977), 'Search theory and duration data : A theory of sorts', *Quarterly Journal of Economics*, 91, 39-57.

Schwarz, G. (1978), 'Estimating the dimension of a problem', *Annals of Statistics*, 6, 461-464.

Sider, H. (1985), 'Unemployment duration and incidence: 1968-82', *American Economic Review*, 75, 461-472.

Spinnewyn, F. (1982), 'De statistische analyse van gegevens in verband met de werkloosheid', Working Paper, Centrum voor Economische Studien, Katholieke Universiteit Leuven.

van den Berg, G.J. (2001), 'Duration models: Specification, identification, and multiple durations', in J.J. Heckman and E. Leamer (eds.), *The Handbook of Econometrics*, Volume V, Amsterdam: North Holland.

van den Berg, G.J., A. Gijsbert, C. van Lomwel, and J.C. van Ours (1998), 'Unemployment dynamics and age', Discussion Paper n°9897, Center for Economic Research.

van den Berg, G.J. and B. van der Klaauw (2001), 'Combining micro and macro unemployment duration data', *Journal of Econometrics*, 102, 271-309.

van den Berg, G.J. and J.C. van Ours (1994), 'Unemployment dynamics and duration dependence in France, The Netherlands and The United Kingdom', *The Economic Journal*, 104, 432-443.

van den Berg, G.J. and J.C. van Ours (1999), 'Duration dependence and heterogeneity in French youth unemployment duration', *Journal of Population Economics*, 12, 273-285.

Van Haeperen, B. (2000), 'Formes d'emploi et flexibilité du marché du travail: Evolutions récentes en Belgique', *Etude spéciale du Service d'Analyse Economique IRES-UCL*, Juin, Louvain-la-Neuve : IRES, Université catholique de Louvain, 117-136.

Appendix

Appendix 1: Empirical autocorrelation coefficient

In this appendix, we derive the empirical correlation coefficients between ε_{kl} and, respectively, ε_{k-1l+1} and ε_{k-1l} .

Let us define the total errors v_{kl} as the sum of the approximation errors and the specification (or measurement) errors:

$$v_{kl} = \omega_{kl} + \varepsilon_{kl} \quad (\text{a1})$$

and make the following assumptions on the covariance structure of these errors:

$$\begin{aligned} E(\omega_{kl}\omega_{k-sl+s}) &= E(\omega_{kl}\omega_{k-sl}) = E(\omega_{kl}\varepsilon_{kl}) \\ &= E(\omega_{k-sl+s}\varepsilon_{kl}) = E(\omega_{k-sl}\varepsilon_{kl}) = 0 \quad \forall s \neq 0 \end{aligned} \quad (\text{b1})$$

First, consider that in (a1), ε_{kl} follows an AR(1) process over k and for $l+k$ fixed:

$$\varepsilon_{kl} = \rho_{l+k}\varepsilon_{k-1l+1} + \zeta_{kl}, \text{ where } |\rho_{l+k}| < 1 \quad (\text{c1})$$

where it is assumed that:

$$E(\varepsilon_{k-1l+1}\zeta_{kl}) = 0 \quad (\text{d1})$$

Upon substitution of (c1) in (a1) and expressing ε_{k-1l+1} as a function of v_{kl} and ω_{kl} , we obtain:

$$v_{kl} = \rho_{l+k}v_{k-1l+1} + \zeta_{kl} + \omega_{kl} - \rho_{l+k}\omega_{k-1l+1} \quad (\text{e1})$$

Multiplying both sides by v_{k-1l+1} and taking the expectation give:

$$\begin{aligned} E(v_{kl}v_{k-1l+1}) &= \rho_{l+k}E(v_{k-1l+1}^2) + E(\zeta_{kl}v_{k-1l+1}) + E(\omega_{kl}v_{k-1l+1}) \\ &\quad - \rho_{l+k}[E(\omega_{k-1l+1}\varepsilon_{k-1l+1}) + E(\omega_{k-1l+1}^2)] \end{aligned} \quad (\text{f1})$$

By using the assumptions stated in (b1) and in (d1), we then find that the covariance of v_{kl} in the dimension $l+k$ and for varying k is given by:

$$E(v_{kl}v_{k-1l+1}) = \rho_{l+k}s_\varepsilon^2 \quad (\text{g1})$$

From (g1), we can finally derive a consistent estimate of ρ_{l+k} :

$$\hat{\rho}_{l+k} = \frac{\frac{1}{N} \sum_{l=0}^{16} \sum_{k=2}^{18-l} \hat{v}_{kl}\hat{v}_{k-1l+1}}{\hat{s}_\varepsilon^2} \quad (\text{h1})$$

where \hat{v}_{kl} is the OLS residual of the regression (13) and \hat{s}_ε^2 is the estimate of s_ε^2 given in (c2) in Appendix 2. In fact, $\hat{\rho}_{l+k}$ is a generalisation of the empirical autocorrelation coefficient found in the literature to the case of two errors terms, one of them following an AR(1) process. The LM-test statistics for $\hat{\rho}_{l+k}$ takes the form (see Godfrey 1978 and Breusch and Pagan 1980):

$$LM = N (\hat{\rho}_{l+k})^2 \underset{\alpha}{\sim}^{H_0} \chi^2(1) \quad (\text{i1})$$

where the null hypothesis is $H_0 : \rho_{l+k} = 0$ against $H_1 : \rho_{l+k} \neq 0$ and α is the significance level of the test.

If in (a1), ε_{kl} follows an AR(1) process over k and in the dimension l :

$$\varepsilon_{kl} = \rho_l \varepsilon_{k-1l} + \zeta_{kl}, \text{ where } |\rho_l| < 1 \quad (\text{j1})$$

where it is assumed that:

$$E(\varepsilon_{k-1l} \zeta_{kl}) = 0 \quad (\text{k1})$$

we can similarly find a consistent estimate of $\hat{\rho}_l$:

$$\hat{\rho}_l = \frac{\frac{1}{N} \sum_{l=0}^{17} \sum_{k=2}^{18-l} \hat{v}_{kl} \hat{v}_{k-1l}}{\hat{s}_\varepsilon^2} \quad (\text{11})$$

and apply the same LM-test statistic than in (i1).

Appendix 2: Two-step estimation procedure

In this appendix, we give the estimation procedure of the mixture model (13) where the disturbances, ε_{kl} , follow an AR(1) process over k and in the dimension $l + k$.

When ε_{kl} follows an AR(1) process across k and for $l + k$ fixed such that:

$$\varepsilon_{kl} = \rho\varepsilon_{k-1l+1} + \xi_{kl} \quad (\text{a2})$$

the mixture model in (13) becomes a non-linear regression model with heteroskedastic and correlated disturbances. Assuming that $E(\omega_{kl}\varepsilon_{kl}) = 0$, the variance-covariance matrix of the N total disturbances, v_{kl} , is then block-diagonal. The block-matrix, idempotent and of dimension $l + k$ (for simplicity, we denote $l + k$ by c), takes the following form:

$$\begin{bmatrix} s_{1c-1}^2 + s_\varepsilon^2 & \rho s_\varepsilon^2 & \rho^2 s_\varepsilon^2 & \dots & \rho^{c-1} s_\varepsilon^2 \\ \rho s_\varepsilon^2 & s_{2c-2}^2 + s_\varepsilon^2 & \rho s_\varepsilon^2 & \dots & \rho^{c-2} s_\varepsilon^2 \\ \rho^2 s_\varepsilon^2 & \rho s_\varepsilon^2 & s_{3c-3}^2 + s_\varepsilon^2 & \dots & \rho^{c-3} s_\varepsilon^2 \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{c-1} s_\varepsilon^2 & \rho^{c-2} s_\varepsilon^2 & \rho^{c-3} s_\varepsilon^2 & \dots & s_{c0}^2 + s_\varepsilon^2 \end{bmatrix} \quad (\text{b2})$$

(c,c)

The estimation procedure consists then in two-steps. The first step consists in using OLS to estimate the residuals \hat{v}_{kl} . These allow us to construct a consistent estimate of the specification errors, \hat{s}_ε^2 :

$$\hat{s}_\varepsilon^2 = \frac{1}{N} \left\{ \sum_{l=0}^{17} \sum_{k=1}^{18-l} \left[(\hat{v}_{kl})^2 - \hat{s}_{kl}^2 \right] \right\} \quad (\text{c2})$$

and to calculate the empirical correlation coefficient ρ (see (h1) in Appendix 1). In a second step, we estimate the variance-covariance matrix in (b2) and estimate the parameters of the mixture model (13) by GLS.

Appendix 3: Baker and Melino procedure

In this appendix, we give the estimation procedure of the mixture model (16) where the heterogeneity distribution is specified as discrete with unknown support and an unknown number of mass points. We follow the estimation method proposed by Heckman and Singer (1984) as adjusted by Baker and Melino (2000). To make their procedure applicable, we need to modify the objective function, i.e. we minimise the weighted sum of squared residuals (*WSSR*) rather than maximise a loglikelihood function.

The approach to identify the optimal number of mass points proceeds as follows (see Baker and Melino 2000, pp.360-361). Suppose that we have the Weighted Least Square (WLS) estimator $\{\hat{\theta}, \hat{v}_s, \hat{\pi}_s; s = 1, 2, \dots, S\}$, where θ are the other parameters, conditional on having S points of support. The approach consists in checking whether the *WSSR* can be decreased by adding an additional point of support to the S existing ones. Let $v_{S+1} = \bar{v}$ be a candidate for such an additional mass point. Formally, the problem is to minimise the *WSSR* over $(\pi_1, \pi_2, \dots, \pi_{S+1})$, given $\{\hat{\theta}, \hat{v}_s; s = 1, 2, \dots, S\}$ and $v_{S+1} = \bar{v}$ (where $\varphi_4^*(l) = \exp[\varphi_4(l)]$):

$$\min_{\{\pi_1, \dots, \pi_{S+1}\}} WSSR = \sum_{l=0}^{17} \sum_{k=1}^{18-l} \left[\frac{1}{\hat{s}_{kl}^2 + \hat{s}_e^2} \left(\left(1 - \hat{P}_{kl} \right) - \frac{\sum_{s=1}^{S+1} \pi_s \exp \left[-v_s \varphi_4^*(l) \sum_{j=1}^k h_{jl} \right]}{\sum_{s=1}^{S+1} \pi_s \exp \left[-v_s \varphi_4^*(l) \sum_{j=1}^{k-1} h_{jl} \right]} \right)^2 \right] \\ + \lambda \left(\sum_{s=1}^{S+1} \pi_s - 1 \right) + \delta \left(\sum_{s=1}^{S+1} \pi_s v_s - 1 \right) + \sum_{s=1}^{S+1} \mu_s \pi_s$$

for which the Kuhn-Tucker first-order conditions (FOCs) are:

$$\frac{dWSSR}{d\pi_s} = A + \lambda + \mu_s = 0 \quad (\text{a3})$$

$$\frac{dWSSR}{d\lambda} = \left(\sum_{s=1}^{S+1} \pi_s - 1 \right) = 0 \quad (\text{b3})$$

$$\frac{dWSSR}{d\delta} = \left(\sum_{s=1}^{S+1} \pi_s v_s - 1 \right) = 0 \quad (\text{c3})$$

$$\text{and } \pi_s \geq 0; \mu_s \geq 0; \mu_s \pi_s = 0 \quad (\text{d3})$$

$$\text{where } A = \sum_{l=0}^{17} \sum_{k=1}^{18-l} \frac{2}{\hat{s}_{kl}^2 + \hat{s}_e^2} \left(\left(1 - \hat{P}_{kl} \right) - \frac{\sum_{s=1}^{S+1} \pi_s \exp \left[-v_s \varphi_4^*(l) \sum_{j=1}^k h_{jl} \right]}{\sum_{s=1}^{S+1} \pi_s \exp \left[-v_s \varphi_4^*(l) \sum_{j=1}^{k-1} h_{jl} \right]} \right)$$

$$\cdot \left(\frac{\exp \left[-v_s \varphi_4^*(l) \sum_{j=1}^k h_{jl} \right] \sum_{s=1}^{S+1} \pi_s \exp \left[-v_s \varphi_4^*(l) \sum_{j=1}^{k-1} h_{jl} \right] - \exp \left[-v_s \varphi_4^*(l) \sum_{j=1}^{k-1} h_{jl} \right] \sum_{s=1}^{S+1} \pi_s \exp \left[-v_s \varphi_4^*(l) \sum_{j=1}^k h_{jl} \right]}{\left(\sum_{s=1}^{S+1} \pi_s \exp \left[-v_s \varphi_4^*(l) \sum_{j=1}^{k-1} h_{jl} \right] \right)^2} \right)$$

The idea then consists in evaluating the Kuhn-Tucker multiplier of the candidate additional mass point, μ_{S+1} , at $\pi_{S+1} = 0$ and for a grid of values of $\bar{v} \in [v_L, v_H]$, given the parameter estimates of the mixture model with S points of support, i.e. given $\{\hat{\theta}, \hat{v}_s, \hat{\pi}_s; s = 1, 2, \dots, S\}$. If $\mu_{S+1} \geq 0$, the last constraint of (d3) is binding, so adding a point of support at \bar{v} cannot decrease the *WSSR*. If this is true for all $\bar{v} \in [v_L, v_H]$, then $\{\hat{\theta}, \hat{v}_s, \hat{\pi}_s; s = 1, 2, \dots, S\}$ is the non-parametric WLS estimator. If instead $\mu_{S+1} < 0$ for some \bar{v} , then it is possible to decrease the *WSSR* by adding a point of support at \bar{v} . This additional mass point will however not be retained if the Schwartz information criterion of the mixture model with $S + 1$ mass points (see (15) in Section 3) is above the one with S mass points (see Baker and Melino 2000, p.376).

If we multiply (a3) by π_s , sum over s , and use (b3), (c3) and (d3), we obtain $\lambda + \delta = 0$. The Kuhn-Tucker multiplier, μ_s , is therefore the (negative of the) derivative of the *WSSR* with respect to π_s :

$$\mu_s = -A \tag{e3}$$

Applying this approach in a systematic way from the outset, the estimation of the non-parametric mixture model (16) can be summarised as follows (see Baker and Melino 2000, pp.361-362):

Step 0: Set $S = 1$ and $\pi_1 = 1$. Choose initial value for θ and v_1 .

Step 1: Given the current value of S ,

minimise the Schwartz Information Criterion (SIC),

i.e. minimise the *WSSR*, over θ and $(v_s, \pi_s; s = 1, 2, \dots, S)$.

a. If $\text{SIC}_S > \text{SIC}_{S-1}$, then **STOP**.

b. Else, **Step 2**.

Step 2: Evaluate μ_{S+1} for a grid value of $\bar{v} \in [v_L, v_H]$ such that P_{kl} is restricted to the set $[\varepsilon, 1 - \varepsilon]$ where ε is set to 10^{-5} .

a. If $\mu_{S+1} \geq 0 \quad \forall \bar{v} \in [v_L, v_H]$, then **STOP**.

b. Else, set v_{S+1} to the value of \bar{v} that yield the smallest value for μ_{S+1} .

Step 3: Given the current value of θ , $(v_s; s = 1, 2, \dots, S)$ and v_{S+1} , increase the value of S by 1 and minimise the *WSSR* to obtain new initial values for $(\pi_s; s = 1, 2, \dots, S)$. Return to **Step 1**.

Tables and Figures

Table 1: Hazard model estimates - Men aged 29-44

Variables	GAMMA		POINTS OF SUPPORT	
		SD		SD
c	-0.59	0.08	-0.71	0.06
1. Duration (in quarters)				
2 ($\gamma_2 - c$)	-0.07	0.03	-0.15	0.03
3 ($\gamma_3 - c$)	-0.05	0.03	-0.16	0.05
4 ($\gamma_4 - c$)	-0.04	0.04	-0.15	0.07
5 ($\gamma_5 - c$)	-0.02	0.05	-0.11	0.08
6 ($\gamma_6 - c$)	-0.07	0.06	-0.16	0.10
7 ($\gamma_7 - c$)	-0.16	0.07	-0.26	0.11
8 ($\gamma_8 - c$)	-0.21	0.08	-0.34	0.12
9 ($\gamma_9 - c$)	-0.15	0.08	-0.32	0.12
10 ($\gamma_{10} - c$)	-0.27	0.10	-0.49	0.14
11 ($\gamma_{11} - c$)	-0.18	0.10	-0.46	0.14
12 ($\gamma_{12} - c$) [*]	-0.31	0.11	-0.70	0.15
2. Calendar time at outflow				
Cycle				
β_{c1}	-0.22	0.05	-0.22	0.05
β_{c2}	0.06	0.02	0.06	0.02
β_{c3}	0.03	0.02	0.02	0.02
β_{c4}	0.04	0.02	0.04	0.02
β_{c5}	0.05	0.02	0.04	0.02
Seasons				
Winter (β_{s2})	-0.31	0.03	-0.30	0.03
Spring (β_{s3})	0.01	0.03	0.02	0.03
Summer (β_{s4})	-0.31	0.03	-0.31	0.03
3. Interaction term				
ϕ	1.99	0.81	0.87	0.35
4. Variance of the Gamma at $l=0$				
δ_0^*	0.12	0.06		
$\sigma_0^2 = \exp(-\delta_0^*)$	0.89	0.05		
4'. Points of support				
π			0.55	0.05
v_1			1.60	0.09
$v_2 = (1 - \pi v_1) / (1 - \pi)$			0.26	0.05
$\text{var}[v] = \pi(v_1 - 1)^2 / (1 - \pi)$			0.44	0.05
5. Calendar time at inflow				
Cycle				
η_c	-0.18	0.18	-0.11	0.15
Seasons				
Autumn (η_{s2})	-0.09	0.03	-0.07	0.03
Winter (η_{s3})	-0.04	0.03	-0.03	0.03
Spring (η_{s4})	-0.09	0.03	-0.07	0.03
6. Empirical correlation coefficients of ϵ_{kl}				
ρ_{1+k}	0.65		0.70	
ρ_1	0.03		0.01	
7. Standard deviation of ϵ_{kl}				
s_e	0.009		0.009	
Number of observations	171		171	
Number of estimated parameters	26		27	
Weighted sum of squared residuals	170.44		171.20	
P-value of the goodness-of-fit test	0.073		0.06	
Jeffrey-Bayes posterior-probability statistics	304.12		310.02	

* $\gamma_{18} = \dots = \gamma_{12}$

Table 2: Hazard model estimates - Men aged ≤ 28

Variables	GAMMA		GAMMA Constant Baseline MPH		POINTS OF SUPPORT Constant Baseline MPH	
		SD		SD		SD
c	-0.45	0.07	-0.49	0.06	-0.71	0.05
1. Duration (in quarters)						
2 $(\gamma_2 - c)$	0.00	0.04				
3 $(\gamma_3 - c)$	0.05	0.08				
4 $(\gamma_4 - c)$	0.04	0.12				
5 $(\gamma_5 - c)$	0.10	0.15				
6 $(\gamma_6 - c)$	0.14	0.18				
7 $(\gamma_7 - c)$	0.11	0.20				
8 $(\gamma_8 - c)$	0.09	0.22				
9 $(\gamma_9 - c)$	0.06	0.24				
10 $(\gamma_{10} - c)$	0.08	0.26				
11 $(\gamma_{11} - c)$	0.14	0.28				
12 $(\gamma_{12} - c)^*$	0.12	0.31				
2. Calendar time at outflow Cycle						
β_{c1}	-0.34	0.04	-0.34	0.03	-0.40	0.03
β_{c2}	0.03	0.02	0.03	0.02	0.04	0.02
β_{c3}	0.03	0.02	0.03	0.02	0.02	0.02
β_{c4}	0.04	0.02	0.04	0.02	0.05	0.02
β_{c5}	0.07	0.02	0.06	0.02	0.05	0.02
Seasons						
Winter (β_{s2})	-0.33	0.04	-0.32	0.04	-0.31	0.04
Spring (β_{s3})	0.01	0.03	0.02	0.03	0.02	0.03
Summer (β_{s4})	-0.26	0.04	-0.26	0.04	-0.26	0.04
3. Interaction term						
ϕ	-0.55	1.61				
4. Variance of the Gamma at $l=0$						
δ^*_0	0.45	0.20	0.59	0.04		
$\sigma^2_0 = \exp(-\delta^*_0)$	1.57	0.314	1.80	0.072		
4'. Points of support						
π					0.67	0.02
v_1					1.36	0.02
$v_2 = (1 - \pi v_1) / (1 - \pi)$					0.29	0.03
$\text{var}[v] = \pi(v_1 - 1)^2 / (1 - \pi)$					0.26	0.05
5. Calendar time at inflow Cycle						
η_c	-0.01	0.19	0.03	0.18	0.47	0.17
Seasons						
Autumn (η_{s2})	-0.08	0.03	-0.08	0.03	-0.06	0.03
Winter (η_{s3})	-0.08	0.03	-0.08	0.03	-0.08	0.03
Spring (η_{s4})	-0.14	0.03	-0.13	0.03	-0.11	0.03
6. Empirical correlation coefficients of ϵ_{ki}						
ρ_{1+k}	0.75		0.74		0.59	
ρ_1	0.04		0.05		0.14	
7. Standard deviation of ϵ_{ki}						
s_ϵ	0.012		0.013		0.016	
Number of observations	171		171		171	
Number of estimated parameters	26		14		15	
Weighted sum of squared residuals	174.18		174.27		158.72	
P-value of the goodness-of-fit test	0.03		0.16		0.42	
Jeffrey-Bayes posterior-probability statistics	307.82		246.25		235.84	

* $\gamma_{18} = \dots = \gamma_{12}$

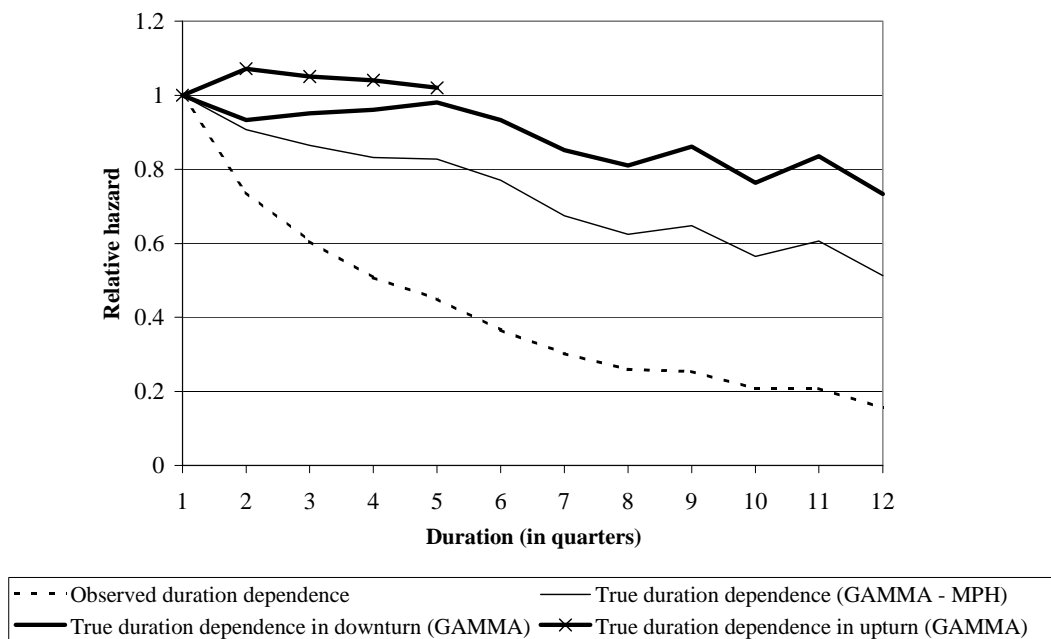


Figure 1: Baseline hazard - Men aged 29-44

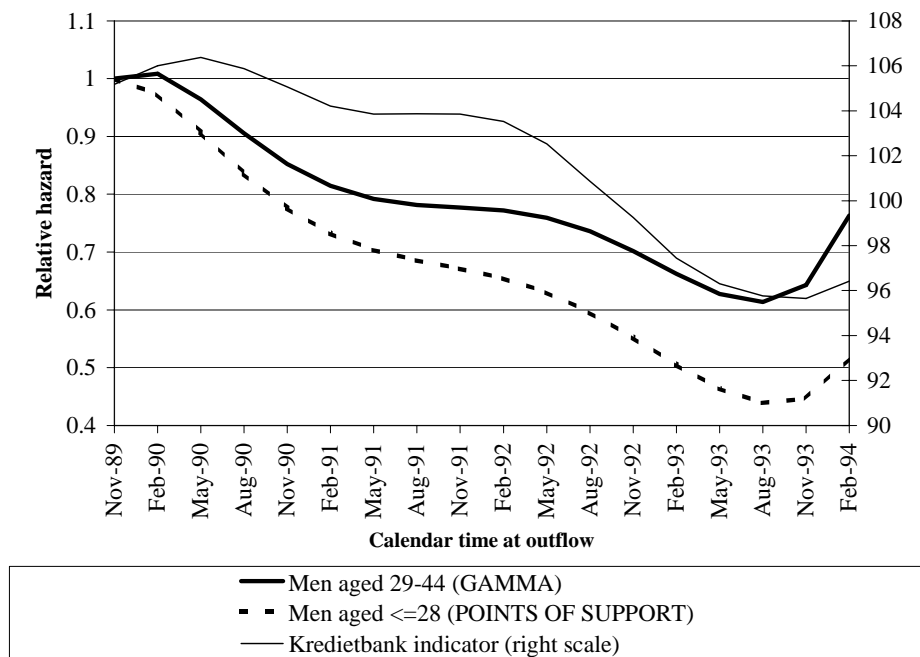


Figure 2: Cyclical dependence of the conditional hazard and the Kredietbank indicator

IZA Discussion Papers

No.	Author(s)	Title	Area	Date
617	L. Magee M. R. Veall	Allocating Awards Across Noncomparable Categories	1	10/02
618	A. L. Booth M. Francesconi G. Zoega	Oligopsony, Institutions and the Efficiency of General Training	6	10/02
619	H. Antecol D. A. Cobb-Clark	The Changing Nature of Employment-Related Sexual Harassment: Evidence from the U.S. Federal Government (1978–1994)	5	10/02
620	D. A. Cobb-Clark	Public Policy and the Labor Market Adjustment of New Immigrants to Australia	1	10/02
621	G. Saint-Paul	On Market Forces and Human Evolution	5	11/02
622	J. Hassler J. V. Rodriguez Mora	Should UI Benefits Really Fall Over Time?	3	11/02
623	A. R. Cardoso P. Ferreira	The Dynamics of Job Creation and Destruction for University Graduates: Why a Rising Unemployment Rate Can Be Misleading	1	11/02
624	J. Wagner R. Sternberg	Personal and Regional Determinants of Entrepreneurial Activities: Empirical Evidence from the <i>REM Germany</i>	1	11/02
625	F. Galindo-Rueda	Endogenous Wage and Capital Dispersion, On-the-Job Search and the Matching Technology	3	11/02
626	A. Kunze	Gender Differences in Entry Wages and Early Career Wages	5	11/02
627	J. Boone J. C. van Ours	Cyclical Fluctuations in Workplace Accidents	5	11/02
628	R. Breunig D. A. Cobb-Clark Y. Dunlop M. Terrill	Assisting the Long-Term Unemployed: Results from a Randomized Trial	6	11/02
629	I. N. Gang K. Sen M.-S. Yun	Caste, Ethnicity and Poverty in Rural India	2	11/02
630	W. Arulampalam	State Dependence in Unemployment Incidence: Evidence for British Men Revisited	3	11/02
631	H. O. Duleep M. Regets	The Elusive Concept of Immigrant Quality: Evidence from 1970-1990	1	11/02
632	B. Cockx M. Dejemeppe	Duration Dependence in the Exit Rate out of Unemployment in Belgium: Is It True or Spurious?	2	11/02