The Impact of Unemployment Duration on Wages: Evidence from French Panel Data 1984-2001

Julien Pouget*, Benoît Rapoport and Salvatore Serravalle*****

May 2010

Abstract

We consider the impact of unemployment duration on wages, using a large-scale longitudinal sample of French workers providing us with very accurate information on earnings and unemployment durations.

Our panel data model potentially suffers from sample selection, unobserved heterogeneity and endogeneity of the unemployment duration: in an attempt to account for all these problems in one framework, we consider the recent estimation method proposed by Semykina and Wooldridge (2005).

For both males and females and for all cohorts, we find consistent evidence of a substantial negative effect of unemployment duration on wages. Unobserved heterogeneity and selectivity cause a downward bias, while endogeneity causes an upward bias in the unemployment duration coefficient.

Keywords: Unemployment, Wages, France, Panel Data, Fixed Effects, Sample Selection, Endogeneity, Instrumental Variables

JEL Classification: J30, J64

---

* DREES, CREST-INSEE, and IZA. E-mail : pouget@ensae.fr

** DREES and University Paris 1-Panthéon-Sorbonne. E-mail: benoit.rapoport@sante.gouv.fr

*** DREES and ENSAE. E-mail: salvatore.serravalle@ensae.fr

We thank participants at the 14th International Conference on Panel Data (Xiamen University) and at the 3S seminar of the Drees. All remaining errors are ours. The views expressed herein are attributable to the authors and do not reflect those of the institutions to which they belong.

1. Introduction

An increasing number of empirical studies suggest that an individual's work experience is not fully represented by the number of years in employment. Two individuals with the same number of years worked may still differ in the frequency of career interruptions. According to the human capital theory, employment breaks are likely to generate mainly negative wage effects. But little is known about the different wage effects of different types of interruption such as unemployment, illness, childbearing, or parental leave.

The aim of this paper is to shed light on the effects of an unemployment spell on the subsequent wage. Not only unemployment spells reduce current earnings, because benefits, when existing, are almost always lower than the wage of the preceding job, but unemployment spells have also an impact on earnings following re-employment. Indeed, in the first place, during unemployment, work experience is not accumulated (Mincer and Ofek, 1993). Moreover, displaced or fired workers loose firm-specific human capital, which, depending on firms and sectors, may be important, even though some (fast) recovery is possible, specially if displaced workers can re-invest into specific human capital (Topel, 1991). Next, decreasing or time-limited unemployment benefits may also lower re-employment earnings through a reduction in the reservation wage. Individuals seem indeed to be more sensitive to benefit exhaustion than to human capital decrease (Belzil, 1995). Unemployment may also be viewed as a stigma that carries out information on the displaced worker and discourages firms to re-employ him, or at least leads them to offer lower wages to displaced workers. Conversely, an unemployment spell might enhance the matching between the worker and his next employer (see Pissarides, 1994).

Most of the studies questioning the causal effect of unemployment on wages have investigated the incidence of an interruption, either on re-employment or on mid/long-term earnings. Most of them have shown that unemployment has a persistent negative effect on earnings (see for example, Ruhm, 1991; one exception is Burda and Mertens, 2001, who obtain on German data a not very high (negative) average effect, which is in fact positive for individuals of the first quartile – but largely negative for workers of the three upper quartiles). This effect is partly due to additional displacement (Stevens, 1997), in particular because a spell of unemployment increases the likelihood of future unemployment spells. Arulampalam (2001) sheds the light on such a scar effect of unemployment on re-employment wage. Only a few recent studies address the question of the effect of the duration of the unemployment spell. , although Gregory and Jukes (2001) show on British data that unemployment duration has a permanent impact on subsequent wages, a one-year spell adding a further wage penalty of 10 percentage points. Incidence however has only a temporary effect, the earning setback being largely eroded over two years after re-employment

Our study brings new empirical evidence on the effect of the unemployment duration on subsequent wages. We use a unique dataset, crated by matching data from private and public sector payroll records (Déclarations Annuelles de Données Sociales (DADS), self-employed employment records, and unemployment insurance records. This sample, the *Echantillon Inter-régimes de Cotisants* (EIC) collected since 2001 by the Directorate for Research, Studies, Evaluation and Statistics (DREES) of the French Ministry of Social Affairs, is a large-scale longitudinal dataset which gives very accurate information on individual path on the labour market. In particular, it contains very precise data on unemployment (duration and reasons). It thus offers new opportunities to investigate the effects of job interruptions, and more specifically of unemployment spells, and subsequent wages, but also on state dependence. Indeed, until now, French panel data did not allow to fully describe individual careers on the labour market. The EIC makes it possible: in particular, we can clearly discern between all reasons explaining why a previously wage earner in the private sector, has little or no wage a given year (whether he became inactive, unemployed, or whether he moved to another sector). As noted by Gregg and Tominey (2005) most of the surveys previously used are rather short: people are followed only for a few years (for example, Gregory and Jukes (2001) follows British men between 1984 and 1994; Arulampalam (2001) between 1991 and 1997; Ruhm (1991) uses the 1971-1975 PSID). Gregg and Tominey (2005) uses the British National Child Development Survey, which allows them to follow individual between ages 23 and 42; however information is not collected on a yearly base, but about every 10 years, so that it suffers from memory bias. The EIC allows us to follow people from 1945 for participation, employment sector and earnings, and from 1984 for unemployment (in fact 1974, but unemployment without benefit is not consistently collected before 1984). As it is a match of administrative records, our dataset does not suffer from memory bias.

We thus focus on the impact of the last unemployment duration on subsequent wages for both men and women. We use several cohorts (born in 1950, 1954, 1958 and 1962) to evaluate what we call hereafter the "return" to unemployment (referring to the so-called return to education based on Mincer equations). Dealing with four cohorts allows controlling for heterogeneity with regard to the macroeconomic environment m. Indeed, labour market situation appears to be different for people who have entered to it in 1970 (people born in 1950) and those who have entered to it 12 years after. Mass unemployment problems have probably been deeper for young generations than for the elder.

Most previous studies have dealt with the problem of unobserved heterogeneity by taking advantage of panel data (for example Stevens, 1997 and the references below). Some of them have taken into account the selection problem (Gregory and Jukes, 2001; Arulampalam, 2001 for example), whereas some others have taken into account the possibly endogenous character of the unemployment duration (Gregg and Tominey (2005). But none of them have dealt with all of those problems.

Indeed, our large scale dataset allows us to simultaneously correct for sample selection, and to take into account unobserved heterogeneity and endogeneity of the unemployment duration. To this aim, we compare two sets of estimators.First, we calculate the new two-step estimator proposed by Semykina and Wooldridge (2005), who show how to estimate panel data models in the presence of selection when the equation of interest contains endogenous explanatory variables. Second, we compute an extension of the semiparametric estimator of Kyriazidou (Charlier *et alii*, 2001). This estimation strategy coupled with other basic estimation strategies allows also to highlight the potential biases of estimators that would not control for any source of bias. The direction of the biasesalso leads us to infer some characteristics of the French labour market performance.

The paper proceeds as follows. Section 2 briefly reviews the predictions from economic theory and the main results of previous empirical studies. The data are described in section 3. Econometric strategy is discussed in section 4, and section 5 presents results.

2. Theory and related research

The effects of unemployment on future employment and earnings have been the subject of many empirical studies. Until recently most of them focused on the consequences of unemployment spells, less on the effects of duration of unemployment. We have learnt – mainly from the United States – that unemployment is followed by a lower path for future earnings after new employment. Indeed, interruptions in the career obviously involve a loss in the current income all along the unemployment spell. But these interruptions also bring a long-term “scar” on individuals who therefore have lower earnings in employment.

We can mention at least two reasons for scarring (negative correlation between unemployment duration and subsequent wages) to occur. A number of models of earnings determination suggest a negative link between the occurrence and duration of unemployment and subsequent wages. First, within the human capital theory, during unemployment, work experience is not accumulated and skills previously accumulated tend to depreciate (Mincer and Ofek, 1993). The worker accumulates firm-specific skills which are rewarded through the wage rate. As theses skills are not transferable from one firm to another, unemployment incidence may lead to permanent wage losses. In addition, the duration of the unemployment spell may have a negative impact on transferable skills. Secondly, the employee’s unemployment history can be regarded as a negative signal (a stigma) by the employer, who has imperfect information about the potential employee’s productivity. The employer can therefore apply a wage penalty that will never be totally overtaken.

More generally, the question of displaced workers has been widely discussed in the economic literature (see Farber, 1999, for a review). One of the main results is that workers who lose involuntarily their job have longer unemployment spells and suffer substantial wage losses

(Addison and Portugal, 1989; Jacobson, Lalonde and Sullivan, 1993). Similar evidence is presented by Rhum (1991). More recently, Gregory and Jukes (2001) estimate the impact of unemployment on earnings following re-employment for a sample of British men. They only find a temporary effect of unemployment incidence, largely eroding after two years; while the effect of unemployment duration is found to be more permanent. Mroz and Savage (2006) estimate the long-term effects of youth unemployment on later labour market outcomes. Using the American National Longitudinal Survey of Youth (NLSY), they find a large and persistent negative effect of prior unemployment on earnings.

Nevertheless, theoretically, job interruption does not necessarily imply a loss in the subsequent earnings. Unemployment is part of the ongoing process of workforce reallocation: some jobs are destroyed, others are created and employees and employers look for the best matches. In this framework, an unemployment spell can be voluntarily chosen by an employee in order to find a better match and subsequent greater wages. Here, unemployment duration would reflect search for a suitable new job and, as long as costs of the additional period of unemployment does not exceed the expected benefit of the new match, the longer the duration, the greater would be the subsequent wages.

This paper addresses the issue of the scarring effect of unemployment in the French case. More precisely we attempt to evaluate the "return" to unemployment duration using panel data of several birth cohorts (born in 1950, 1954, 1958 and 1962) on the period 1984-2001. Data include workers with unemployment spells, those with no interruption to their employment experience and those who never participate to the labour market. This allows us to control for heterogeneity and to identify the effects of unemployment after taking into account both observed and unobserved characteristics of those individuals who do, and those who do not, experience unemployment.

From an econometric point of view, our paper tackles the same problems as Semykina and Wooldridge (2005) who show how to estimate panel data models in the presence of selection when the primary equation contains endogenous explanatory variables. He is also similar to Jäckle (2007), who investigates the effects of health on wages with panel data estimates considering selection (the decision to participate in the labour market is non-random) and endogeneity of health status (as the self-reported health variable could induce measurement error and omitted variable bias). In our paper, we consider the same selection process (i.e. to be employed or not) and the endogeneity of duration of unemployment (as this variable – as previously shown – may cause a simultaneity and reverse causality bias). In order to address the question of robustness, we also present results derived from an extension of the semi-parametric estimator of Kyriazidou (1997).

3. Data

Our sample is derived from the *Echantillon Inter-régimes de Cotisants* (EIC hereinafter), a new large-scale longitudinal sample of French workers, maintained by the Directorate for Research, Studies, Evaluation and Statistics (Drees) of the French Ministry of Health and Solidarity in order to provide information about retirement rights along the lifecycle. This sample is based upon matched administrative data sets, providing us with very accurate information on earnings, unemployment spells and personal and job characteristics.

The EIC dataset links various longitudinal data. The main data source is the "Déclarations Annuelles de Données Sociales" (DADS), an administrative database of matched employer-employee information collected by Insee (Institut National de la Statistique et des Etudes Economiques). These data are based upon mandatory employer reports of the gross earnings and the annual number of days in the private sector for each employee subjects to French payroll taxes (Abowd, Kramarz and Margolis, 1999). We calculate real daily wage rates from the DADS (in €-2002): this individual real wage rate (taken in logarithm) is the variable of interest in our study. For people who have been employed in several firms in the same year, and as we work on annual data, wages and employment duration are aggregated and we retain firm and job characteristics corresponding to the highest annual wage received that year by the employee.

The EIC links the DADS with other administrative data maintained by Unedic, the French administration in charge of unemployment insurance. This dataset contains precise information about unemployment spells : their duration is measured in days and the reasons of each spell are detailed at a very thin level. We derive from these data the duration of the last spell of unemployment and will investigate its effect on wages. Before 1984, data are incomplete because they do not contain unemployment spells of people who did not receive government compensations. Since 1984, Unedic data are complete, including unemployment periods without government compensation. That's the main reason of the choice to start our study in 1984. However, pre-84 career is taken into account when we compute experience and past events that help to identify labour market participation. Note that data about unemployment duration are also right-censored since we observe individuals only until year 2001. This will not be specifically treated in this paper but we must keep it in mind when interpreting the results.[1] The data provide us with the reasons of each unemployment spell. We have aggregated these reasons into five groups: layoff for economic reason, layoff for personal reasons, resignation, end of a contract, and other reasons. We make the hypothesis that these reasons have a non-direct impact on wage path but that the impact passes through its effect on unemployment duration. This hypothesis will be statistically tested. These reasons will be used as instruments in our econometric estimations, in order to correct for endogeneity of unemployment duration in the wage equation.

As far as we know, this is the first study using the DADS linked with the Unedic dataset. The added-value of this match is obvious, because we are now able to distinguish unemployment spells when workers are not observable in the DADS panel. Our initial sample, observed between 1984 and 2001, covers workers alive in 2001 and born between 1934 and 1970 (with one cohort every four years). In each cohort, 2.67% of individuals are sampled. As a result, the longitudinal sample contains data for nearly 100,000 workers. For practical reasons, in this study, we only present results for cohorts born in 1950, 1954, 1958 and 1962. Since our study aims to estimate the "return" to unemployment, we decided to limit our sample to workers who made the main part of their career in the private sector (people who really face unemployment risk) or were out of the labour market. More precisely, our sample excludes

[1] This is not a real problem as we will assume that the potential effects of the duration of employment are linear.

workers who have obtained more than 16 quarters of retirement right thanks to another activity than working in the private sector from 1984 to 2001. Thus, people whose main activity occurred in the agricultural sector, as an independent or in the public sector are excluded from the sample.

In order to avoid measurements problems in the DADS concerning the number of days worked (which is fundamental in the derivation of the wage rate, calculated per day), we decided to consider that people who worked less than 15 days during a given year were in fact inactive during this year. This allows avoiding very high wage rates that would not correspond to a real situation. People who had no activity during the period are included in our sample so that they will play a role only in the selection equation but not in the wage equation. This point is important to note, because it means that, in our study, unemployment and inactivity status will be considered identically on the selection process in the sense that wage is not available in these two cases. Additionally, we also decided to exclude people whose wage rate is above the 199th half-percentile of the wage distribution. This is done each year so that even if an individual is out the “window selection” for only one year, this individual will not be in our final balanced sample. Exclusion of people whose wage rate is above the 199th half-percentile concerns about 3% of the whole sample. We decided to keep the first half-percentile of the wage distribution, because exclusion of concerned people would lead to drop about another 5% to 6% of the sample, which is quite large (similarly, dropping the first and the last percentile of the wage distribution leads to drop 12% to 14% of each cohort). In order to test the robustness of those choices, all the estimations have also been conducted after excluding people whose at least one wage was under the first percentile or above the 99th percentile of the wage distribution. The results are very similar to those we present here; however, in this case, the estimated effects of unemployment duration on wage and the estimated biases are slightly weaker than those presented. Note also that people who have at least one very low wage have more probably encountered an unemployment spell, so that dropping those people might be a source of bias.

At this step, our dataset does not provide us with the diploma. We will thus not control its effect directly but through unobservable heterogeneity. However, as soon as the diploma does not vary for an individual during the whole period, the Semykina-Wooldridge procedure does not allow to separately identify its effect from the individual fixed effect. Another point we want to mention is the absence, in our data, of the number of children and of the marital status. This would have played a major role in the selection equation, particularly for women, especially because those variables may vary across the period.

In the following, it is worth keeping in mind that cohort comparisons we make here mix up different effects. The 4 cohorts are observed during the same period (1984-2001), and thus at different ages. Moreover, they have experienced mass unemployment at different ages.

Table 1 presents the sample characteristics about unemployment. The most recent generations face a more important unemployment risk than former generations. Indeed, whereas 66.5% of men born in 1950 have never been unemployed until 2001, 47.0% of men born in 1962 have been in this case. More generally, recent generations tend to be more frequently unemployed than former generations. This is consistent with the fact that unemployment affects more frequently young people than the elder. In our sample, women tend to be less subjected to unemployment than men: a larger proportion of them have never been unemployed from the beginning of the career to 2001. Even if in France, labour participation of women has substantially risen (from 55.3% in 1984 to 61.8% in 2001), this reflects the fact that women still participate less than men (the proportion is stable, around 75.0%).[2] Killingsworth and Heckman (1986) provide a formal analysis and a discussion on this topic.

[Insert Table 1]

Figure 1 clearly shows how unemployment spells can affect the wage path. The more frequently individuals have been unemployed, the lower are their wages. Whereas in 1984, when individuals are 26 years old, at the beginning of the career, wages are quite close, in 2001, individuals who have never been unemployed have wages 50% higher than individuals who have been unemployed more than three times. Note the slope of the curves in the middle of the 90's. This is mainly due to treatment changes that have involved a break in the DADS series in 1993.

[Insert Figure 1]

Regarding unemployment duration (figures 2 and 3), there is no clear disparity between generations except for the cohort born in 1962 which seems to face short unemployment more frequently than the other generations. This reflects the fact that, simultaneously with the rise of unemployment, young people have to face a lot of short unemployment spells whereas elder individuals face less unemployment spells of larger duration. For men, about one third of unemployment spells last less than 3 months. 75% of these spells last less than one year. For women, unemployment durations tend to be greater : about two thirds of women spell last less than one year.

[2] Data from Labour Force Survey made annually by OECD.

[**Insert Figure 2]**

[**Insert Figure 3]**

It is interesting to compare wage path conditionally of the time spent in the last unemployment spell (figure 4), making no distinction between the dates when these spells occurred. At the beginning of the career, there in no clear disadvantage between individuals who have never been unemployed and those whose duration of last spell of unemployment is less than one year. But the gap is increasing along the career and for the generation born in 1958, at 40 years old, the wage rate of workers never unemployed is about 24% higher than the wage rate of individuals who have been unemployed at least once and whose last spell is less than one year. For workers whose last spell of unemployment is more than one year, the gap is large from the beginning of the career.

[**Insert Figure 4]**

[**Insert Table 2]**

Table 2 deals with the causes of unemployment. End on contract is the main reason of being unemployed: between 40% and 50% of unemployment spells are due to an end of contract. It is more likely the case for young people who, in France, are more often offered fixed duration contract than their elder (Junod, 2006). On the contrary, older generations are more concerned by layoffs for economic reason. Note that resignation is twice more important for women than for men. One explanation could be that many women tend to resign after a maternity leave or to follow their spouse when he gets a new job.

The effects of those different causes on unemployment duration appear to be heterogeneous. This duration is the most important for individuals who have been laid off because of non economic reasons. An unemployment spell due to an end of contract is likely to be shorter than for other causes. But in return, the wage impact seems to be deeper (see Figure 5). This can be explained by the fact that, until 2001, end of contracts are likely to concern more low qualified individuals: their wage would tend to be structurally lower than other individuals who faced unemployment spells for other reasons. Note that unemployment duration for women is always longer than for men (about 3 months more, in average).

[Insert Figure 5]

The wages path of individuals who have been unemployed is a curve which reaches its maximum in 1993. This is partly due to changes in the treatment of the DADS production. There are also other explanations for that scope. Indeed, there is evidence that, since the middle of the 90's, French workers' wages have slowed down (see Desplatz, Jamet, Passeron and Romans, 2003). On one part, that's may due to the monetary policy of the early 90's which aimed to slow down inflation in preparation for the entry in the eurozone; on the other part, the reduction in work time (35-hour work week instead of 39 in France since 1998) has been accompanied by salary moderation (often through wage freezes over a year or two). Globally, there is no big difference in wage path depending on the cause of unemployment, maybe except for "end of contract". This may mean that causes of unemployment do not directly cause the wage path but has only an impact through the differential duration they involve. Practically, this will justify our choice to use causes of unemployment as instruments for unemployment duration in the empirical implementation.

4. Empirical implementation

As panel data become widely used in empirical economics, a number of recent papers provide a range of solutions to reduce the bias due to selectivity, but also endogeneity. First, many data sets are unbalanced panels : a number of studies, surveyed by Baltagi and Song (2006) have addressed the problem of selectivity, but under the assumption of strictly exogenous explanatory variables. The most popular are Wooldridge (1995) and Kyriazidou (1997). Wooldridge's procedure is very similar to Heckman's two-stage estimator; it allows the unobserved fixed effects and the covariates to be correlated in both the selection and primary equations. The semi-parametric estimator developed by Kyriazidou (1997) relies on pairwise differences over time for individuals having the same "selection" characteristics. Rochina-Barrachina (1999) proposes a third estimator based, as Kyriazidou's, on pairwise differences but involving more parameterization than Kyriazidou's. Dustmann and Rochina-Barrachina (2007) propose an extensive discussion of these three methods and, in particular, a nice comparison of underlying assumptions.

Extensions to allow for endogenous explanatory variables in the primary equation are more recent: Vella and Verbeek (1999), and, more recently, Fernández-Val I. and Vella (2007) provide some guidelines. In our study, we use the new method proposed by Semykina and Woodridge (2005) in order to take account of unobservable heterogeneity, selection bias, and endogeneity of one of the explanatory variable (i.e. the duration of the last unemployment spell). Duration of the last unemployment spell is likely to be endogenous through its correlation with unobserved heterogeneity, associated with individual ability and motivation. The most motivated and skilled individuals should find a new job faster and have lower subsequent unemployment duration. We also implement the Kyriazidou estimator. Charlier *et alii* (2001) show that it can be quite straightforwardly extended to take into account endogenous regressors (see also Askildsen *et alii*, 2002).

We remind below the theoretical econometric framework proposed by Semykina and Wooldridge (2005) to estimate panel data models when a panel is unbalanced due to selection, and some explanatory variables are endogenous. In all that follows, parameters indexed by 1 refer to the wage equation; while parameters indexed by 2 are related to the selection equation.

Main equation of interest (wage equation)

Let y_{it1} be the wage of individual *i* in year *t*. Then, for *i* = 1, …, *N* and *t* = 1, …, *T*, the wage equation can be written as:

$$y_{it1} = x_{it1}\beta_1 + c_{i1} + u_{it1}$$

where:

- x_{it1} is a $1 \times K$ vector of explanatory variables, which contains both exogenous and endogenous variables;
- β_1 is a $K \times 1$ vector of explanatory parameters;
- c_{i1} is an individual unobserved effect;
- and u_{it1} is the idiosyncratic error term.

In addition, let $z_{it} = (z_{it}^{(1)} \vdots z_{it}^{(2)})$ be a $1 \times L\ (L \geq K)$ vector of instruments, which are sufficiently correlated with the explanatory variables, and strictly exogenous, conditionally on c_{i1}. We consider two types of instruments. On one part, there are instruments which allow to correct for endogeneity (noted $z_{it}^{(1)}$) ; on another part there are instruments which help to identify the selection process (noted $z_{it}^{(2)}$). In the current paper of Semykina and Wooldridge (2005), this distinction does not appear but here, it will help for the understanding of our approach. $z_{it}^{(2)}$ must be always observed whereas $z_{it}^{(1)}$ must be observed at least when the selection indicator is unity. $z_{it}^{(1)}$ includes all the exogenous explanatory variables in x_{it1}. We note $z_i^{(j)} = (z_{i1}^{(j)},..., z_{iT}^{(j)})$ for j = 1, 2.

Selection equation

In all that follows, s_{it2} is the selection indicator; i.e. (y_{it1}, x_{it1}) are assumed to be observed when $s_{it2} = 1$. This selection indicator is generated by the following latent variable

$$s_{it2}^* = z_{it}^{(2)}\delta_2 + c_{i2} + u_{it2}$$

so that $s_{it2} = \mathbf{1}\{s_{it2}^* > 0\}$, and where:

- $z_{it}^{(2)}$ is the vector of instruments described above;
- δ_2 is a vector of parameters;
- c_{i1} is an individual unobserved effect;
- and u_{it2} is the idiosyncratic error term.

Following Semykina and Wooldridge (2005), we have to make the following assumptions:

i) $u_{it2} \big| z_i^{(2)}, c_{i2} \sim N(0,1)$, so that s_{it2} follows an unobserved effects probit model.

ii) The unobserved effect can be modeled as $c_{i2} = \eta_2 + \bar{z}_i^{(2)}\xi_2 + a_{i2}$ (following Mundlak, 1978); where $a_{i2} \big| z_i^{(2)} \sim N(0, \tau_2^2), \quad t = 1,\dots,T$

In this general case, the selection indicator can be written as

$$s_{it2} = \mathbf{1}\{s_{it2}^* > 0\} = \mathbf{1}\{\eta_2 + z_{it}^{(2)}\delta_2 + \bar{z}_i^{(2)}\xi_2 + a_{i2} + u_{it2} > 0\} = \mathbf{1}\{\eta_2 + z_{it}^{(2)}\delta_2 + \bar{z}_i^{(2)}\xi_2 + v_{it2} > 0\}$$

where $v_{it2} | z_i^{(2)} \sim N(0, 1+\tau_2^2), \quad t = 1, \ldots, T$.

In fact, in our particular study, the coefficients in assumption ii) will not be restricted to be the same at the different periods, so that selection is defined by:

$$s_{it2} = \mathbf{1}\{\eta_{t2} + z_{it}^{(2)}\delta_{t2} + \bar{z}_i^{(2)}\xi_{t2} + v_{it2} > 0\}$$

where $v_{it2} | z_i^{(2)} \sim N(0,1), \quad t = 1, \ldots, T$

Testing for selection bias (Procedure A hereinafter)

Under the above assumptions, Semykina and Wooldridge (2005) derive a test for selection bias.

1) For each time period, we estimate the probability of selection by using a probit model:

$$\mathrm{P}(s_{it2} = 1 | z_i^{(2)}) = \Phi(\eta_{t2} + z_{it}^{(2)}\delta_{t2} + \bar{z}_i^{(2)}\xi_{t2})$$

2) We compute the estimated inverse Mills ratios

$$\hat{\lambda}_{it2} \equiv \lambda(\hat{\eta}_{t2} + z_{it}^{(2)}\hat{\delta}_{t2} + \bar{z}_i^{(2)}\hat{\xi}_{t2})$$

3) We estimate the augmented main equation by using fixed-effect-2SLS (FE-2SLS). The main equation is augmented by adding the interactions of the inverse Mills ratios with time dummies :

$$y_{it1} = x_{it1}\beta_1 + c_{i1} + \sum_{\tau=1,\ldots,T}(\rho_{\tau 1}\hat{\lambda}_{it2}\mathbf{1}\{t = \tau\}) + e_{it1} \qquad i = 1,\ldots, N \quad t = 1,\ldots, T$$

4) Last, we test the joint significativity of the $\rho_{\tau 1}$'s by using the Wald test (or alternatively the Student test if the pooled probit model is used for the fist step). As

suggested by Semykina and Wooldridge (2005), we use a variance matrix robust to serial correlation and heteroskedasticity (see Wooldridge, 2002).

In the case where selection has been detected (ie in the case where the Wald (or the Student) test rejects the null assumption of no-selection (H_0: $\rho_{\tau 1}=0$, $\tau=1,...,T$)), Semykina and Wooldridge (2005) propose the following method to correct for selection bias.

Correcting for selection bias (Procedure B hereinafter)

Semykina and Wooldridge (2005) make the additional assumptions, besides (i) and (ii):

iii) $\mathrm{E}(u_{it1}|z_i, v_{it2}) = \mathrm{E}(u_{it1}|v_{it2}) = \rho_{t1}v_{it2}, \quad t=1,\ldots,T$

iv) $c_{i1} = \eta_1 + \bar{z}_i\xi_1 + a_{i1}, \quad \text{where} \mathrm{E}(a_{i1}|z_i, v_{it2}) = \mathrm{E}(a_{i1}|v_{it2}) = \phi_{t1}v_{it2}, \quad t=1,\ldots,T$

Under assumptions (i) to (iv), the following procedure corrects for selection bias:

1) For each time period, we estimate the probability of selection by using a probit model:

$$\mathrm{P}(s_{it2}=1|z_i^{(2)}) = \Phi(\eta_{t2} + z_i^{(2)}\delta_{t2} + \bar{z}_i^{(2)}\xi_{t2})$$

2) We compute the estimated inverse Mills ratios

$$\hat{\lambda}_{it2} \equiv \lambda(\hat{\eta}_{t2} + z_{it}^{(2)}\hat{\delta}_{t2} + \bar{z}_i^{(2)}\hat{\xi}_{t2})$$

3) We estimate the augmented main equation by using pooled-2SLS using $(z_{it}, \bar{z}_i, \hat{\lambda}_{it2})$ as instruments. The main equation is augmented by adding the interactions of the inverse Mills ratios with time dummies, and the $\bar{z}_i^{(1)}$:

$$y_{it1} = x_{it1}\beta_1 + \eta_1 + \bar{z}_i\xi_1 + \sum_{\tau=1,\ldots,T}(\gamma_{\tau 1}\hat{\lambda}_{it2}\mathbf{1}\{t=\tau\}) + e_{it1} \quad i=1,\ldots,N \quad t=1,\ldots,T$$

4) We correct the variance matrix as described by Semykina and Wooldridge (2005).

The variance matrix must be corrected because of the well-known problem of generated regressors (Murphy and Topel, 1985). In the second step, estimators of inverse Mills-ratios are used instead of the Mills ratios which are not observed.

By using matrix notations the pooled 2SLS estimator of $\theta = (\eta_1, \beta_1^{'}, \xi_1^{'}, \gamma_1, ..., \gamma_T)$ on the selected sample is:

$$\hat{\theta} = ((W'H)(H'H)^{-1}(W'H)')^{-1}(W'H)(H'H)^{-1}H'y$$

where W is the matrix of regressors containing $\hat{w}_{it} = (1, x_{it1}, \bar{z}_i, 0, \ldots, \hat{\lambda}_{it2}, 0, \ldots, 0)$, and H is the matrix of instruments containing $h_{it} = (1, z_{it1}, \bar{z}_i, 0, \ldots, \hat{\lambda}_{it2}, 0, \ldots, 0)$.

Thus

$$\sqrt{N}(\hat{\theta} - \theta) = \left(\left(\frac{W'H}{N} \right) \left(\frac{H'H}{N} \right)^{-1} \left(\frac{W'H}{N} \right)' \right)^{-1} \left(\frac{W'H}{N} \right) \left(\frac{H'H}{N} \right)^{-1} \left(\frac{H'((W - \hat{W})\theta + E_1)}{\sqrt{N}} \right)$$

where E_1 is the second-step residual.

Only the last term requires some work to compute. Semykina and Wooldridge (2005) show that it can be estimated by

$$\frac{1}{N} \sum_i \hat{p}_i \hat{p}_i^{'} \text{, where } \hat{p}_i = \sum_t s_{it2} h_{it}^{'} \hat{e}_{it1} - \hat{F}\hat{r}_i \,.$$

We have

$$\hat{F} = \frac{1}{N} \sum_i \sum_t s_{it2} h_{it}^{'} (\hat{\theta}' \nabla_{\pi} \hat{w}_{it}^{'}),$$

where $\nabla_{\pi} \hat{w}_{it}^{'}$ is the Jacobian of $\hat{w}_{it}^{'}$ with respect to the first step vector of parameters $\pi = (\pi_1', ..., \pi_T')$, with $\pi_t = (\eta_{t2}, \delta_{t2}', \xi_{t2}')'$.

The parameter $\hat{r}_i$ is computed by stacking the $\hat{r}_{it}$, where

$$\hat{r}_{it} = (-E(H_t))^{-1} \frac{\phi(q_{it}\hat{\pi}_t)}{\Phi(q_{it}\hat{\pi}_t)(1-\Phi(q_{it}\hat{\pi}_t))} q_{it}^{'}(s_{it2} - \Phi(q_{it}\hat{\pi}_t)) \,,$$

In this expression, q_{it} is the vector of variables, $\hat{\pi}_t$ the estimator of the vector of parameters, and $E(H_t)$ the consistent estimator of minus the expected Hessian of the probit for period *t*.

Note that, in procedure A, the first stage estimation of the $\hat{\pi}_t$ does not affect the limited distribution of the *t* statistic under the assumption of no-selection effect (Wooldridge, 1995), so that there is no need to correct the variance matrix, when testing for selection.

Kyriazidou estimator

Under an assumption of "conditional exchangeability" of the error-term which implies that the "selection effect" for a given individual does not depends on the period, Kyriazidou (1997) shows that it is possible in the two-period case to first-differentiate the main equation in order to eliminate both the fixed effect and the selection effect for individuals *i* and pairs of periods (s,t) for which $z_{it}^{(2)}\delta_2 = z_{is}^{(2)}\delta_2$. In practice, for most of individuals, no pair (s,t) with $z_{it}^{(2)}\delta_2 = z_{is}^{(2)}\delta_2$ exists. Kyriazidou thus suggests to differentiate across observations when $z_{it}^{(2)}\delta_2$ and $z_{is}^{(2)}\delta_2$ are close. In practice, she suggests the following procedure:

1. get consistent estimates of the parameters in the selection equation. Then
2. construct "kernel weights" that are decreasing functions of the difference between $z_{it}^{(2)}\delta_2$ and $z_{is}^{(2)}\delta_2$, and
3. get the parameters of the main equation by running a weighted least-square regression.

Charlier *et alii* (2001) show that Kyriazidou's procedure can be extended, first when individuals are observed during more than 2 periods, and, second, when one (or several) regressor(s) is endogenous. In the second case, provided that one has at least one instrument the last step of the procedure can be replaced by a weighted IV regression. In the first case, the two last steps can be applied to each pair of waves, by using individuals who participate the two waves. The parameters of the main equation are then estimated by using the minimum distance estimator, which is a weighted average of the estimators of each pairs (s,t), with the optimal weighting matrix. As shown by Charlier *et alii* (2001), the optimal weights are given by the inverses of the covariance matrix estimates of the different pairs.

As in Askildsen *et alii* (2002), we use in the first-step a conditional logit model on the sample of individuals who change status over time. We also use a Gaussian kernel, and a sequence of bandwidth proportional to $n^{-1/5}$: $h_n = hn^{-1/5}$.

We thus estimate, in the second step, a weighted IV regression with weights $\omega_i = \frac{1}{h_n} K\left(\frac{z_{it}^{(2)}\hat{\delta}_2 - z_{is}^{(2)}\hat{\delta}_2}{h_n}\right)$, where K is the standard normal density, $h_n = hn^{-1/5}$ and $\hat{\delta}_2$ is estimated in the first step.

Choosing *h* is quite uneasy as shown by Kyriazidou, as the final estimator depends on this choice. Kyriazidou proposes a plug-in procedure in order to choose *h*, but she shows that the "final" *h* depends on the initial guess. Empirically, we observed that in all cases (the 4 cohorts * 2 genders = 8 cases) the minimum distance estimator converges when *h* grows. We thus choose *h* accordingly.[3]

In our empirical application, we will report the coefficient estimates for six different estimation methods: pooled OLS, pooled 2SLS, fixed effects OLS, fixed effects 2SLS, Semykina and Wooldridge procedure B, and IV Kyriazidou, in order to be able to distinguish the effects of the 3 potential sources of bias (unobserved heterogeneity, selection, and endogeneity).

[3] *h* is then fixed at 100 (our sample is quite large).

5. Results

In the first step, we need instruments in order to identify the selection process. Those instruments must be always observed. Remember that our selection indicator separates workers with observed wage from individuals which wage is not observed (mainly unemployed or inactive but also in sector which is not, strictly speaking, the private sector). Our instruments are the past elements of the individual's career. Indeed, the EIC[4] panel provides us with number of validated quarters in each French retirement fund (there are more than 30 retirement funds in France). We have aggregated them : quarters validated thanks to a private-sector activity, during unemployment, during sick leave, thanks to an activity in the agricultural sector, to an activity as an independent, to an activity in firms with special system, to an activity in local authority and finally to an activity in the public sector. The lags of the cumulated number of validated quarters are used as instruments. In fact, we assume that the past career is a good predictor of the current career and will capture a part of the unobserved heterogeneity that affects the selection process. Since we have restricted our sample to individuals who have made the main part of the career in the private sector, we are aware of the threshold of our instruments. An individual is considered as a participant if he reports positive wage in a given year.

Tables 5 to 12 in the Appendix present the different estimations of the wage equation on panel data where the duration of the last spell of unemployment is included among explanatory variables and is likely to be endogenous. Moreover, there is a selection process. We will test

[4] For more details, see Caillot L., Chaput H., Colin C., El Mekkaoui de Freitas N., Michaudon H. (2003), "Echantillon interégimes des cotisants (EIC) : procédures statistiques de constitution de l'EIC", Document de travail n°50, séries statistiques, Drees, Ministère des Affaires Sociales.

the selectivity bias and we will correct if necessary, applying procedures A and B described in Section 4. We now present our six estimators.

Pooled OLS estimator assumes that all explanatory variables are uncorrelated with unobserved heterogeneity and are also strictly exogenous. *Pooled 2SLS* estimator instruments for the duration of the last spell of unemployment, but does not remove unobserved effects. *Fixed effects* estimator allows for correlation between the explanatory variables and unobserved heterogeneity while *FE-2SLS* estimator further allows duration of the last spell of unemployment to be correlated with the idiosyncratic errors. Nevertheless, FE-2SLS estimator assumes that selection into the workforce is not systematically related to idiosyncratic changes in the wage equation. *Procedure A* tests for contemporaneous selection bias (results are presented in Table 3 and 13 in Appendix). In fact, this procedure detects if the selection at time t is correlated with the idiosyncratic error in the wage equation at time t. If there is no evidence of selection bias, the FE-2SLS estimator is consistent. If the null hypothesis in the previous tests is rejected, *procedure B* allows estimating a model with consistency. This procedure corrects for contemporaneous selection. The last is the Kyriazidou estimator.

The set of explanatory variables includes duration of the last unemployment spell (information is available on a daily basis and is expressed in years hereafter) plus demographics and occupational variables (sector of activity, living in Paris, employment status,...) to control for individual heterogeneity. We add employment experience which is the number of years worked by the individuals since 1984. Unemployment duration is not considered strictly exogenous while the other variables are. Particularly, we assume strict exogeneity for experience. This hypothesis is quite strong but data do not provide with sufficient information to correct for potential endogeneity of both experience and unemployment duration. We use the following variables as instruments for employment duration: the duration of the previous unemployment spell, but distinguishing the reason of this spell (layoffs for economic reason, resignation, end of a work contract,…). In the 8 cases (4 cohorts, both for men and women), overidentifying restrictions are not rejected at the 1% significance level (Hansen's test).[5]

As women and men's behaviours on the labour market are quite different, we have chosen to calculate estimations by sex. Sector and time dummies are included in each procedure but the associated parameters are not reported. Figures 6 and 7 present the effects unemployment duration on wage for the 6 estimators.

[5] Those tests are not real formal tests, as we are not aware of overidentifying restrictions tests that would correspond to the Semykina and Wooldridge or to the extended Kyriazidou estimators, that is that would take into account selection and unobserved heterogeneity. Hansen's tests have been implemented after the pooled IV models.

[Insert Figure 6]

[Insert Figure 7]

Figures 6 and 7 show that the duration of the last unemployment spell has a negative and significant impact on daily wage rate, for all cohorts, both for men and women, and for all the estimators. On the whole, this effect is higher for men than for women and its absolute value decreases with the cohort, suggesting that women and young worker are less penalized. For men, unemployment is the main cause of selection, due to the high male participation, but this is not the case for women who may also be penalized by no-participation. For young cohorts, one interpretation is that as they *ceteris paribus* receive lower wages, they might be less penalized because of minimum wage, which might implied that they could also be penalized by a lower probability of finding a new job.[6]

The pooled-OLS estimator is affected by three potential sources of bias: unobserved heterogeneity, endogeneity of the unemployment duration (correlation with the idiosyncratic errors, even after we remove the unobserved effect) and sample selection. We will comment the estimators in comparison with this estimation. For all generations, the "return" to unemployment estimated with the pooled OLS method is about –0.13 for men and about -0.08 for women. Not surprisingly, the point estimate is clearly reduced by controlling for unobserved effects but it is still statistically significant. The fixed effect estimator is roughly –0.07 for all generations for men, and between -0.05 and -0.07 for women[7]. When allowing

[6] As already mentioned comparisons between cohorts are not easy.

[7] Our fixed effect estimator is the within estimator. As the Semykina and Wooldridge estimator models unobserved heterogeneity as suggested by Mundlak (1978), we have also computed a fixed effect estimator (and a fixed effect 2SLS estimator) by modelling the individual effect à la Mundlak. In most cases, the estimated effects of the employment duration variable are very close to those presented in Figures 6 and 7. They are

correlation between unemployment duration and errors (pooled 2SLS estimator) or with idiosyncratic errors (FE-2SLS estimator) the estimated "return" to unemployment is, in most cases, slightly higher, but indeed very close. This indicates that the endogeneity bias is slightly positive (see figure 6): this means that an exogenous positive shock on wages is likely to make the unemployment duration longer (endogeneity bias due to inverse causality). This result is consistent with the job search model. All other things being equal, the higher the wages offered on the labour market, the higher the reservation wage and subsequently the longer the unemployment duration. At this step, we have analysed two sources of bias: unobserved heterogeneity and endogeneity. Whereas removing unobserved effects tends to lower the "return" to unemployment, controlling for regressor's endogeneity tends to increase the "return" to unemployment.

We now turn to the third source of bias: sample selection. Table 3 (Table 13 for women) presents the test, made robust to arbitrary serial correlation and heteroskedasticity, for selection bias according to the procedure A –FE-2SLS estimation with inverse Mills ratio terms added in the primary equation. As argued by Semykina and Wooldridge (2005), procedure A is not a consistent correction procedure as it does not rely on "good" assumptions (for example, procedure A assumes independence between unobserved effects in selection equation and wage equation).

[Insert Table 3]

For all generations, there is some evidence of selection bias that we plan to correct either with the procedure B or by the extended Kyriazidiou's estimator, based on, as it has been explained

slightly lower (in absolute value) for men but only for the simple fixed effect estimator (not for the FE-2SLS estimator), which indeed increases the bias due to the omission of unobserved heterogeneity.

earlier, appropriate assumptions. “Return” to unemployment obtained with the procedure B is between –0.07 and -0.09 for men, and between -0.05 and -0.07 for women, and is higher in the FE-2SLS estimation, excepted for men born in 1958 and 1962. A comparison to the Kyriazidou's estimator shows that the return (in absolute value) is higher with the FE-2SLS estimator than with the Kyriazidou's estimator, in all cases excepted for men born in 1950. Correcting for the selection bias thus generally tends to lower the “return”. In most cases (mainly excepted for men born in 1950), the Kyriazidou's estimator results in lower "return" than the Semykina and Wooldridge's estimator. The differences between the two estimators are very small for women and for men born in 1954 (and 1950): less than 1 point of percentage; they are higher for men born in 1958 and 1962.[8]

Thus, individuals who would have been employed would have been less penalised on their wages after an unemployment spell. We can explain this result by some institutional features of the French labour market. Men who are unemployed (or inactive) are likely to be the less qualified and their wages, close to the minimum wage, cannot fall *under* the minimum wage. Thus, correcting for the selection bias reduces “return” to unemployment essentially because wage decreases are limited on the bottom of the wage distribution. For men, this is less clear with the Semykina and Wooldridge's estimator than with the Kyriazidou's estimator, at least for men born in 1958 and 1962, ie the youngest cohorts (selection is often less obvious for men than for women).

Concerning the other explanatory variables, no particular aspect needs to be highlighted, except the role played by experience. In all estimations, the effect of experience is positive and significant, and, as in Semykina and Wooldridge (2005), taking into account unobserved heterogeneity and selection decreases the returns of experience. Additionally, working part-

[8] Note that it is for these two cases that the instruments appear to be the weaker. It may be the case that there is less variability on the previous unemployment spell for the young cohorts due to shorter careers.

time reduces, as expected, the wage, and the effect is lower when selection is taken into account.

Regarding the causes of unemployment spells, we could assume that “return” to unemployment is different depending, for example, on whether workers are fired for economic reasons or whether they resign. Table 4 presents a Procedure B estimation in which unemployment durations of the last spell of unemployment for each cause are simultaneously used as regressors (plus all the other exogenous variables) but without correcting for endogeneity. Indeed, data do not provide with sufficient information to deal with more than one endogenous explanatory variable. Consequently the estimation hereafter only takes into account unobserved heterogeneity and sample selection. Then it does not allow a formal test of the assumption of a differentiate effect by cause: if a more complete data set is made available in the future (with more potential instruments), the question of the different impacts of each cause is worth further thorough investigation.

[Insert Table 4]

The estimations presented in Table 4 do not appear to differ greatly when studying the impact of different causes of unemployment. The “return” to unemployment for workers who were laid off for personal reasons is generally slightly more negative than for other reasons. This may reflect the fact that being laid off for personal reasons is, in incomplete information, a particularly bad signal towards employers. Employers prefer to invest in “faithful” employees and therefore penalise those who seem not to be.

At the opposite, when unemployment is the consequence of an end of contract, one would expect the “return” to its duration to be lower, in absolute value. This might reflect the fact that an end of contract is the most exogenous (and also the easiest to anticipate) cause of unemployment regarding the behaviour of the individuals. When people are hired with a fixed-term contract, there is no ambiguity that subsequent unemployment spell does not rely on their responsibility. When individuals resign or are laid off for personal reasons, according to future employers, the personal responsibility can be engaged and wage penalty is greater. But according to the data, when unemployment is caused by an end of contract, the "return" to its duration is not significantly different from others causes (both for men and women, expected for men born in 1954). This absence of significant dependence on the causes of unemployment tends to validate our initial strategy (i.e. not distinguishing different impact factors for each cause of unemployment and using durations by cause of the previous spell as instruments).

5. Conclusion

In this study we have estimated the causal effect of unemployment duration on the subsequent wages of French men and women born in 1950, 1954, 1958 and 1962 over the period 1984-2001. According to the human capital theory and to the stigma theory, we find consistent evidence of scarring: "return" to unemployment is negative and statically significant. The originality of our study also relies on the data and the estimation methods. Indeed, this is the first study using the DADS linked with uemployment insurance records. This allows to distinguish unemployment spells when workers are not observable in the DADS dataset. Regarding the estimation method, we use the theoretical econometric framework proposed by Semykina and Wooldridge (2005) and an extension of the semiparametric model of Kyriazidou (1997) to estimate panel data models when a panel is unbalanced due to selection, and when some explanatory variables are endogenous. This allows us to correct for three potential sources of bias: unobserved heterogeneity, endogeneity and sample selection. Correcting the unobserved heterogeneity not surprisingly tends to lower the "return" of unemplyment duration on wages. Correcting endogeneity of unemployment duration highlights a result consistent with the job search model: c*eteris paribus*, the higher the wages offered on the labour market, the higher the reservation wage and subsequently the longer the unemployment duration. Finally, correcting for the selection bias tends to lower the "return" essentially because wage decreases are limited on the bottom of the wage distribution (maybe due to some institutional features of the French labour market, e.g. a relative high minimum wage).

This new dataset will allow investigating in the future other aspects of the dynamic of careers and of wage profiles. In particular, a spell of unemployment increases the likelihood of future unemployment spells. As a consequence a first spell of unemployment may function as a "scar" that may have persistent effects 20 years later (Gregg and Tominey, 2005), especially if individuals cannot avoid the repetition of unemployment spells. Stewart (2007) brings also to the fore the linkage between low-wage jobs and repetition of job loss. This speaks in favour of a deeper modelling of the dynamics of the relationships between unemployment and earnings.

*   *
      *

References

Abowd J., Kramarz F. and D. Margolis (1999), "High-Wage Workers and High-Wage Firms," Econometrica, Vol. 67, No, pp. 251-333.

Addison J. T. and P. Portugal (1989), "Job Displacement, Relative Wage Changes, and Duration of Unemployment," *Journal of Labor Economics*, Vol. 7, No. 3, pp. 281-302.

Arulampalam W. (2001), "Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages," *The Economic Journal*, Vol. 111, No. 475, pp. F585-F606.

Askildsen, Jan Erik, Baltagi, Badi H. and Holmas, Tor Helge, (2002), "Will Increased Wages Reduce Shortage of Nurses? A Panel Data Analysis of Nurses' Labour Supply," CESifo Working Paper No 794.

Baltagi B.H., Song S.H. (2006), "Unbalanced Panel Data: A Survey", *Statistical Papers*, Vol. 47, pp. 493-523.

Beffy M., Buchinsky M., Fougère D, Kamionka T. and F. Kramarz (2006), "The Returns to Seniority in France (and Why are they Lower than in the United States?)," CEPR working paper No 5486.

Belzil, C. (1995), "Unemployment duration stigma and re-employment earnings", *Canadian Journal of Economics*, vol. 28(3), pp. 568-85.

Burda, M. C. and Mertens A. (2001), "Estimating wage losses of displaced workers in Germany," *Labour Economics*, Volume 8(1), pp. 15-41.

Caillot L., Chaput H., Colin C., El Mekkaoui de Freitas N., Michaudon H. (2003), "Echantillon interégimes des cotisants (EIC) : procédures statistiques de constitution de l'EIC", Document de travail n°50, séries statistiques, Drees, Ministère des Affaires Sociales.

Charlier, Erwin & Melenberg, Bertrand & van Soest, Arthur (2001), "An analysis of housing expenditure using semiparametric models and panel data," *Journal of Econometrics*, vol. 101(1), pp. 71-107

Desplatz R., Jamet S., Passeron V. and Romans F. (2003), “La modération salariale en France depuis le début des années 1980”, in *Economie et Statistiques* n°367.

Dustmann C., and M.E. Rochina-Barrachina (2007), “Selection correction in panel data models: an application to the estimation of females' wage equations”, Econometrics Journal, vol. 10, pp. 263-293.

Farber, H. S. (1999), “Mobility and Stability: The Dynamics of Job Change in Labor Markets,” in Ashenfelter, O., and David Card (eds.), *Handbook of Labor Economics*, Volume 3, pp. 2439-2483.

Fernández-Val I., Vella F. (2007), “Bias Corrections for Two-Step Fixed Effects Panel Data Estimators”, IZA Working Paper n°2690.

Gregg, P. and Tominey, E. (2005), "The wage scar from male youth unemployment," *Labour Economics*, vol. 12(4), pp. 487-509.

Gregory M. and R. Jukes (2001), “Unemployment and Subsequent Earnings: Estimating Scarring among British Men 1984-94,” *The Economic Journal*, Vol. 111, No. 475, pp. F607-F625.

Jäckle R. (2007), “Health and Wages. Panel Estimates Considering Selection and Endogeneity”, IFO Working Paper No. 43, University of Munich.

Jacobson L. S, LaLonde R. J. and Sullivan D. G. (1993) “Earnings Losses of Displaced Workers,” *American Economic Review*, Vol. 83, No. 4, pp. 685-709.

Junod B. (2006), “Le CDD : un tremplin vers le CDI dans deux tiers des cas…mais pas pour tous” *Document d’études*, DARES.

Killingsworth M.R. and J.J Heckman (1986), “Female labor supply: A survey”, in Ashenfelter O. and R. Layard, editors, *Handbook of Labor Economics*, North-Holland, Amsterdam.

Kyriazidou E. (1997), “Estimation of a Panel Data Sample Selection Model”, *Econometrica*, vol. 65, n°6, pp. 1335-1364.

Mincer J. and H. Ofek (1982), “Interrupted Work Careers: Depreciation and Deterioration of Human Capital”, *The Journal of Human Resources*, Vol. 17, No.1, pp. 3-24.

Mroz T.A and T.H. Savage (2006), “The Long-Term Effects of Youth Unemployment*,” The Journal of Human Resources*, Vol. 41, No.2, pp 259-293.

Murphy K.M and R.H Topel (1985), “Estimation and Inference in Two-Step Econometric Models”, *Journal of Business and Economic Statistics*, vol. 3, n°4, pp 370-379.

Rochina-Barrachina M. E. (1999), “A New Estimator for Panel Data Sample Selection Models”, *Annales d’Economie et de Statistique*, n°55-56, pp. 153-182.

Ruhm C. J. (1991), “Are Workers Permanently Scarred by Job Displacements?,” *American Economic Review*, Vol. 81, No 1, pp. 319-324.

Semykina A. and J. M. Wooldridge (2005), “Estimating Panel Data Models in the Presence of Endogeneity and Selection: Theory and Application,” Working Paper, Michigan State University.

Stevens A. H. (1997), “Persistent Effects of Job Displacement: The Importance of Multiple Job Losses,**”** *Journal of Labor Economics*, Vol. 15, No. 1, Part 1, pp. 165-188.

Stewart M. B. (2007), "The interrelated dynamics of unemployment and low-wage employment," *Journal of Applied Econometrics*, vol. 22(3), pp. 511-531.

Topel, R. (1991), "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority", *Journal of Political Economy*, Vol. 99(1), pp. 145-176.

Vella F., Verbeek M. (1999), “Two-Step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias”, *Journal of Econometrics*, vol. 90, pp. 239-263.

Wooldridge J. M. (1995), "Selection Corrections for Panel Data Models Under Conditional Mean Independence Assumptions," *Journal of Econometrics,* Vol. 68, pp. 115-132.

Wooldridge J. M (2002): Econometric Analysis of Cross Section and Panel Data, MIT Press.

Tables

Table 1 – Sample characteristics about unemployment

	Born in			
Percent of :	1950	1954	1958	1962
Men	*47.6%*	*48.2%*	*48.6%*	*49.7%*
No unemployment spell	66.5%	63.3%	58.4%	47.0%
1 unemployment spells	13.1%	12.8%	13.7%	16.0%
2 unemployment spells	6.8%	8.1%	8.6%	10.9%
3 unemployment spells	4.5%	4.9%	6.0%	7.3%
4 unemployment spells	2.9%	3.3%	3.8%	5.4%
More than 5 unemployment spells	6.2%	7.7%	9.4%	13.4%
Women	*52.4%*	*51.8%*	*51.4%*	*50.3%*
No unemployment spell	69.5%	64.2%	58.2%	47.3%
1 unemployment spells	11.9%	13.9%	14.6%	15.3%
2 unemployment spells	7.2%	8.0%	9.7%	11.6%
3 unemployment spells	3.8%	4.9%	6.1%	8.9%
4 unemployment spells	2.6%	3.1%	3.9%	5.6%
More than 5 unemployment spells	5.0%	5.9%	7.5%	11.4%

Source: EIC2001 (Drees)
Reading: 6.2% of men born in 1950 have been unemployed at least 5 times during the period 1984-2001.

Table 2 – Causes of unemployment

	Men				Women			
	1950	1954	1958	1962	1950	1954	1958	1962
In percentage of total spells								
Layoff for economic reasons	24.4	20.8	19.4	14.5	22.1	20.9	18.2	13.8
Layoff for personal reasons	13.9	14	12.8	11.2	9.9	8.6	7.8	6.8
Resignation	4.5	5.5	4.9	4.8	7.2	8.2	10.1	9.1
End of contract	40.7	43.6	47.9	51.5	42.1	44.6	46.5	53.7
Other reasons	16.5	16.1	15.1	18.1	18.7	17.7	17.5	16.6
Mean duration of unemployment (in days)								
Layoff for economic reasons	340	323	297	281	438	414	388	372
Layoff for personal reasons	453	409	398	371	510	548	524	479
Resignation	307	265	294	223	429	400	407	370
End of contract	281	271	248	212	304	291	294	264
Other reasons	386	378	346	275	481	482	430	380

Source: EIC2001 (Drees)

Reading: for the individuals born in 1950, 24.4% of the unemployment spells are due to a layoff for economic reason.

Sample: at least once unemployed individuals.

Table 3 – Fisher test on the Mills ratio terms - Men

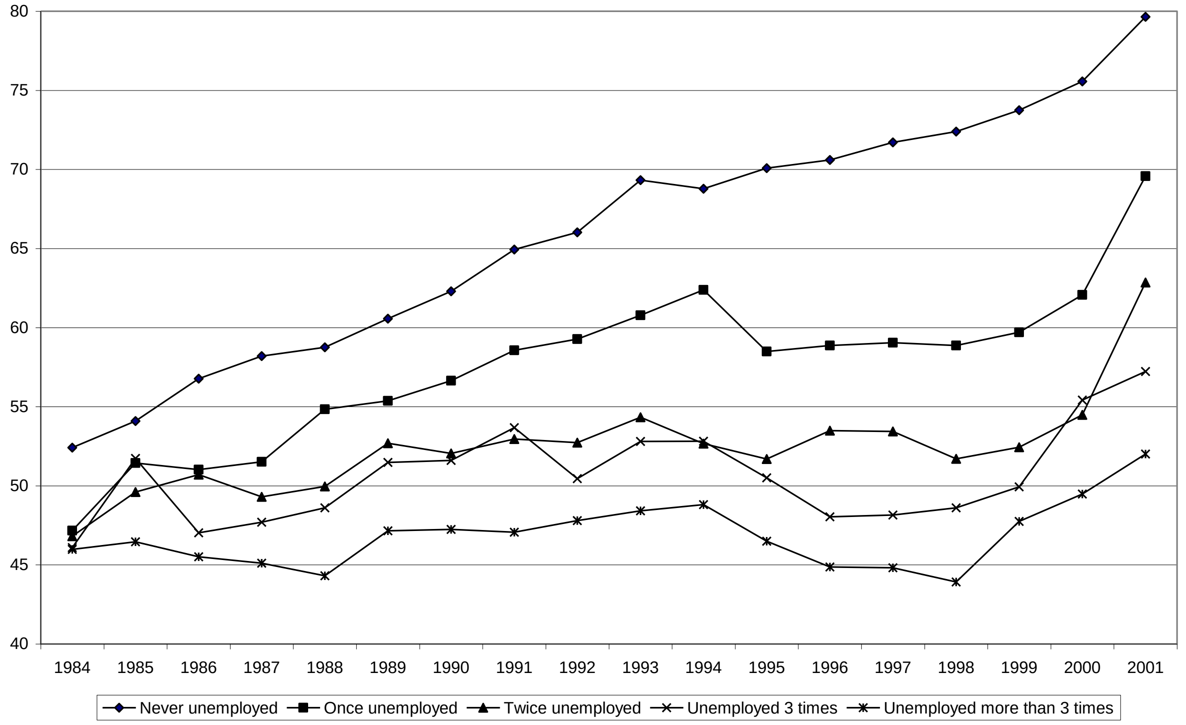
	Generation 50	Generation 54	Generation 58	Generation 62
Fisher statistic	39.1	45.2	61.2	104.9
Significance	<0.001	<0.001	<0.001	<0.001

Table 4 – “Return” to unemployment by cause (Procedure B estimation) - Men

Causes	Generation 50	Generation 54	Generation 58	Generation 62
All causes	-0,052	-0,053	-0,061	-0,050
	(0,007)	(0,007)	(0,006)	(0,006)
Layoff for economic reasons	-0,066	-0,065	-0,068	-0,054
	(0,013)	(0,013)	(0,012)	(0,013)
Layoff for personal reasons	-0,047	-0,068	-0,065	-0,076
	(0,015)	(0,014)	(0,012)	(0,013)
Resignation	-0,085	-0,059	-0,040	-0,070
	(0,035)	(0,028)	(0,024)	(0,031)
End of contract	-0,039	-0,029	-0,049	-0,055
	(0,011)	(0,009)	(0,009)	(0,009)
Other reasons	-0,058	-0,065	-0,077	0,012
	(0,017)	(0,014)	(0,013)	(0,012)

Figures

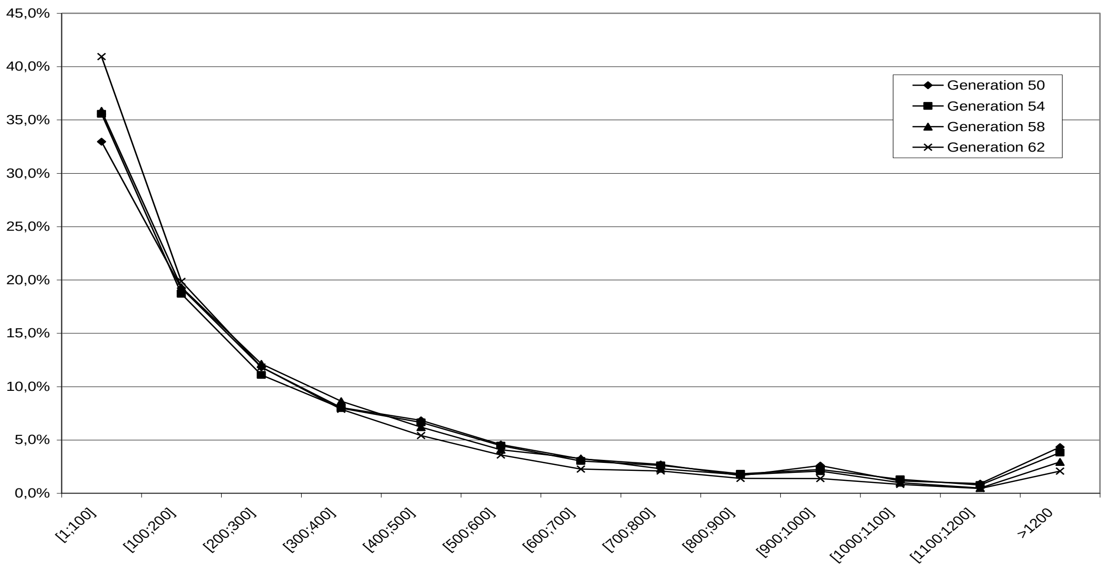
Figure 1 - Unemployment occurrences and wage path – Generation 1958



Source: EIC2001 (Drees)

Reading: in 1995, daily real wage rate of individuals who have been unemployed once during the period 1984-2001 is 58.5 euros.

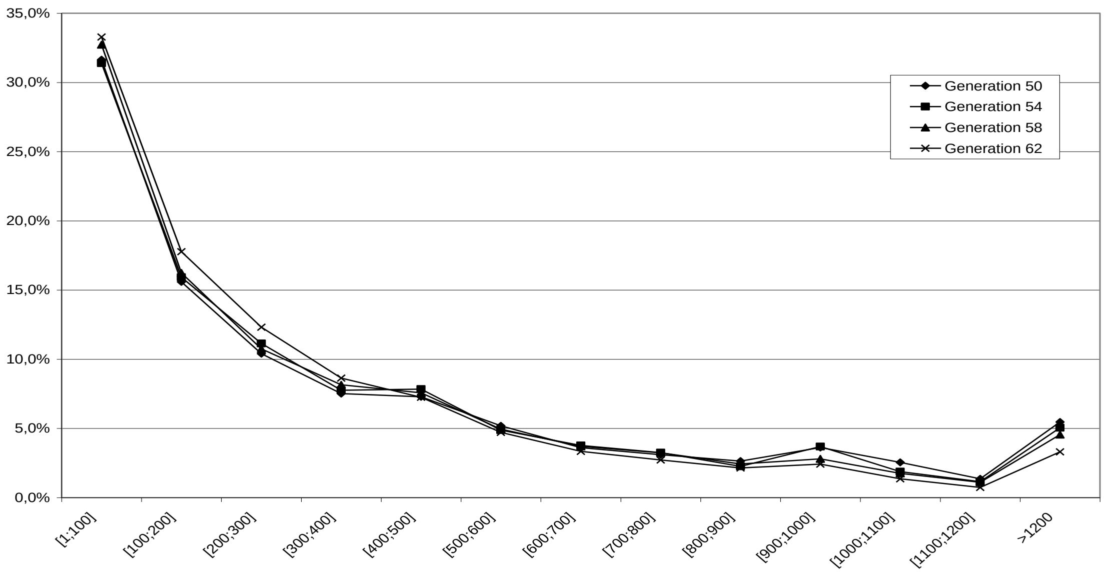
Figure 2 - Distribution of unemployment duration –Men



Source: EIC2001 (Drees)

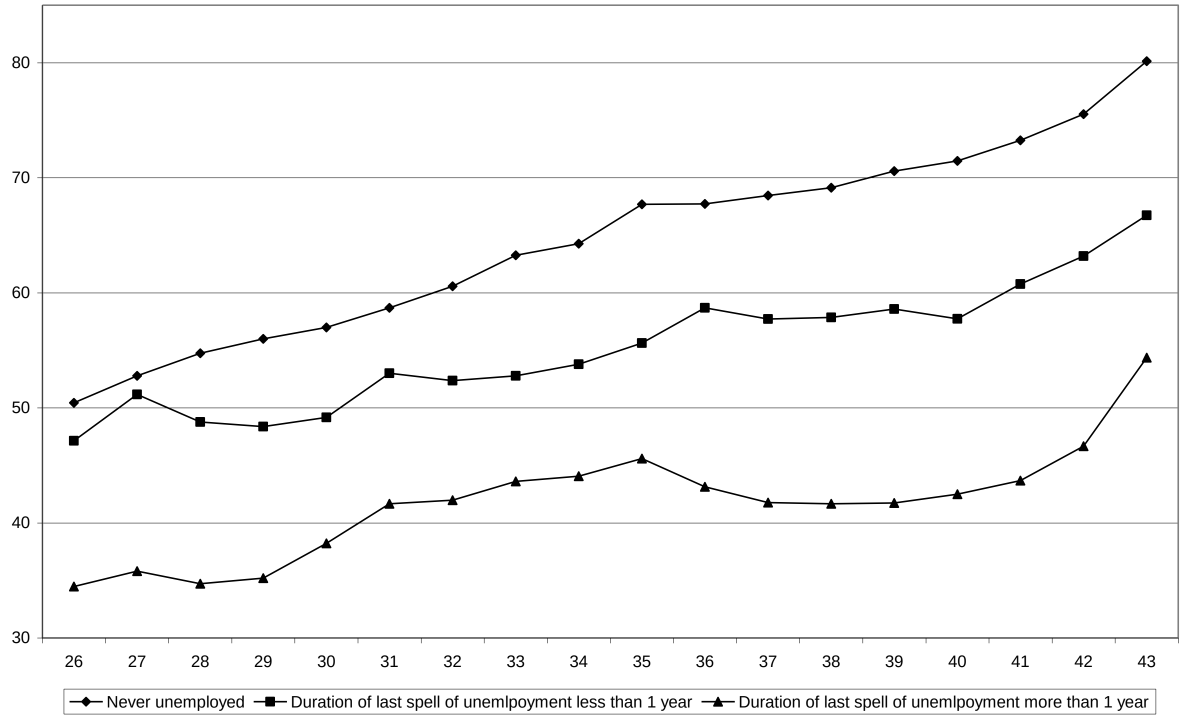
Reading: duration of unemployment has been less than 100 days for 25% of men unemployment spells during the period 1984-2001.

Figure 3 - Distribution of unemployment duration – Women



Source: EIC2001 (Drees)

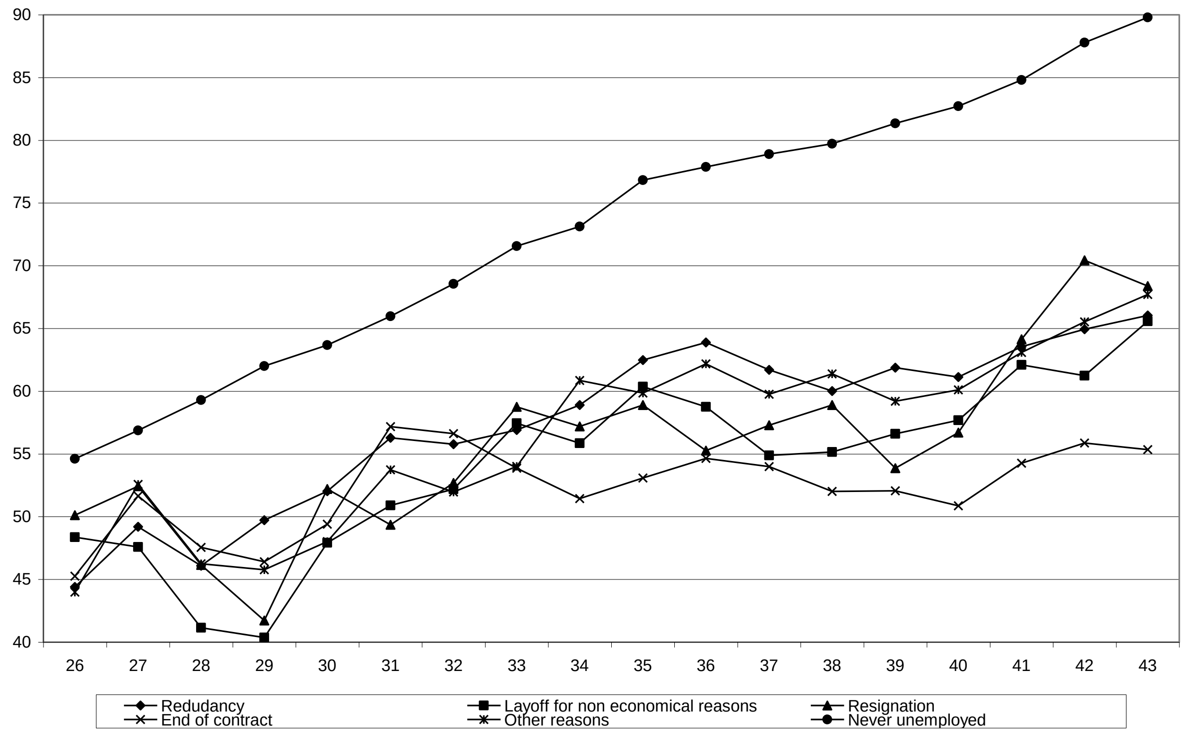
Figure 4 – Wage path for the generation born in 1958 (in €-2002)



Source: EIC2001 (Drees).

Reading: in 1984, the daily real wage rate for people (26 years old) whose last spell of unemployment is less than 1 year is 47.1 euros.

Figure 5 - Real wage rate by cause of unemployment (men born in 1958)



Source: EIC2001 (Drees).

Reading: the real wage rate of individuals, aged 35 years old (in 1993) and who have resigned in the previous years is 58.7 euros.

Figure 6 –“Return” to unemployment – Men

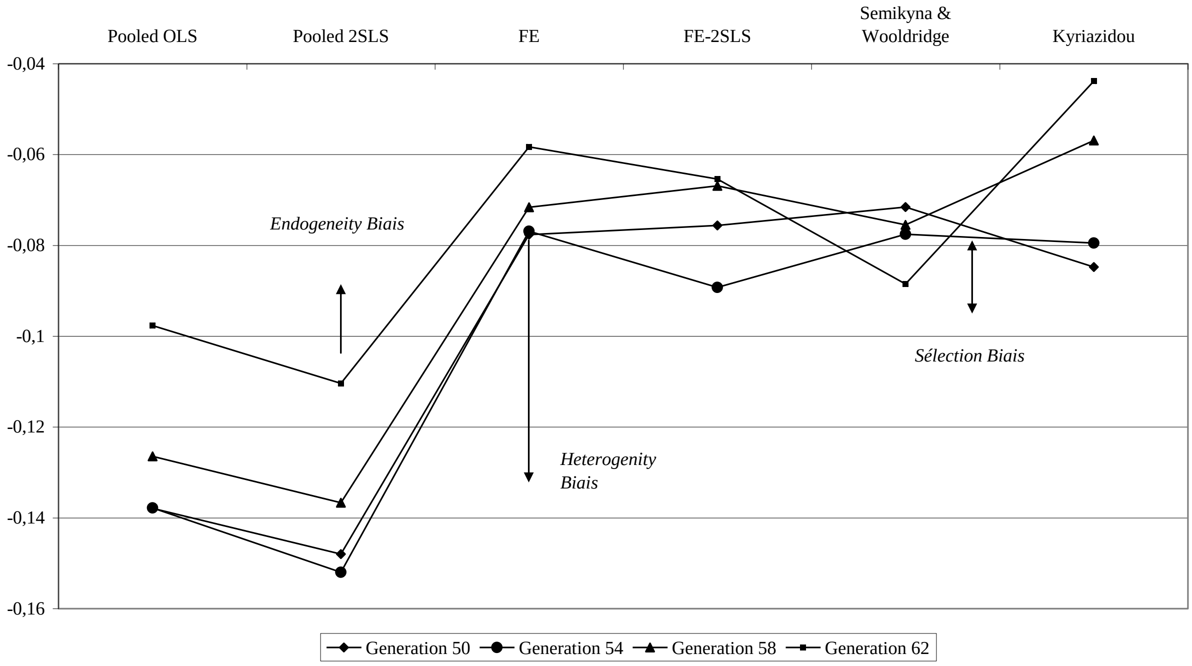
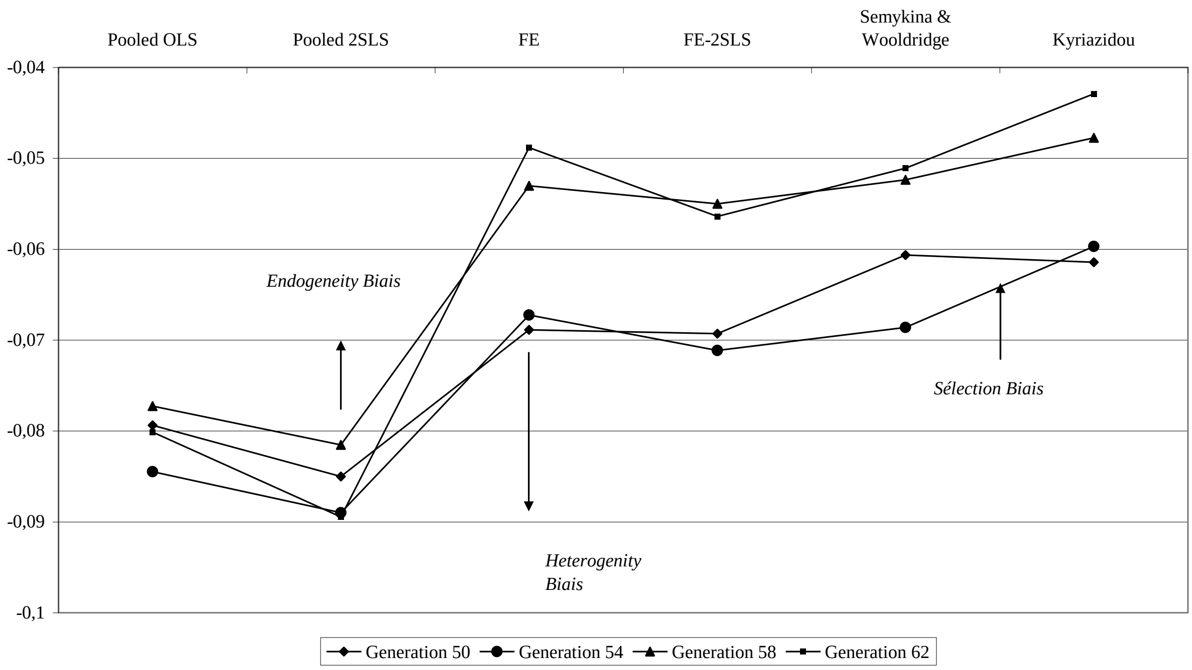


Figure 7 – “Return” to unemployment - Women



Appendix

Table 5 - Estimates for the Log(Real Daily Wages) - Equation for men born in 1950

Explanatory variables	Pooled OLS	Pooled 2SLS	Fixed Effects	FE-2SLS	Procedure B (S&W)	Kyriazidou
Unempl. duration	-0,138	-0,148	-0,078	-0,076	-0,072	-0,085
	(0,007)	(0,009)	(0,003)	(0,004)	(0,01)	(0,009)
Living in Paris	0,162	0,162	0,092	0,092	0,063	0,079
	(0,011)	(0,011)	(0,014)	(0,014)	(0,028)	(0,029)
Part-time	-0,546	-0,542	-0,375	-0,375	-0,325	-0,34
	(0,023)	(0,023)	(0,007)	(0,007)	(0,015)	(0,014)
Experience/10	0,42	0,421	0,302	0,301	0,292	0,341
	(0,03)	(0,03)	(0,015)	(0,015)	(0,026)	(0,035)
Experience²/100	-0,068	-0,068	-0,043	-0,043	-0,037	-0,048
	(0,006)	(0,006)	(0,003)	(0,003)	(0,005)	(0,007)
Manager, salaried of its own firm	0,738	0,736	0,126	0,126	0,107	0,12
	(0,035)	(0,035)	(0,015)	(0,015)	(0,027)	(0,031)
Executive	0,816	0,814	0,151	0,151	0,133	0,136
	(0,014)	(0,014)	(0,009)	(0,009)	(0,014)	(0,019)
Mid-class working	0,387	0,386	0,057	0,058	0,047	0,053
	(0,01)	(0,01)	(0,007)	(0,007)	(0,01)	(0,014)
White collar	0,126	0,125	-0,016	-0,016	-0,023	-0,013
	(0,016)	(0,016)	(0,008)	(0,008)	(0,014)	(0,018)

Sector dummies and time dummies are included in each procedure but not reported.
Standard errors and robust to serial correlation and heteroskedasticity in parentheses under coefficients estimates.

Table 6 - Estimates for the Log(Real Daily Wages) - Equation for men born in 1954

Explanatory variables	Pooled OLS	Pooled 2SLS	Fixed Effects	FE-2SLS	Procedure B (S&W)	Kyriazidou
Unempl. duration	-0,138	-0,152	-0,077	-0,089	-0,078	-0,079
	(0,007)	(0,009)	(0,003)	(0,004)	(0,009)	(0,009)
Living in Paris	0,151	0,151	0,083	0,081	0,063	0,077
	(0,01)	(0,01)	(0,013)	(0,013)	(0,024)	(0,028)
Part-time	-0,515	-0,508	-0,354	-0,352	-0,305	-0,327
	(0,019)	(0,019)	(0,007)	(0,007)	(0,015)	(0,015)
Experience/10	0,507	0,509	0,418	0,425	0,361	0,426
	(0,028)	(0,028)	(0,015)	(0,015)	(0,026)	(0,034)
Experience²/100	-0,094	-0,094	-0,071	-0,073	-0,054	-0,073
	(0,007)	(0,007)	(0,004)	(0,004)	(0,006)	(0,008)
Manager, salaried of its own firm	0,685	0,683	0,154	0,152	0,136	0,143
	(0,04)	(0,04)	(0,017)	(0,017)	(0,03)	(0,036)
Executive	0,834	0,832	0,184	0,182	0,167	0,172
	(0,014)	(0,014)	(0,01)	(0,01)	(0,015)	(0,021)
Mid-class working	0,387	0,385	0,071	0,07	0,06	0,067
	(0,009)	(0,009)	(0,007)	(0,007)	(0,01)	(0,015)
White collar	0,129	0,128	-0,011	-0,011	-0,013	-0,006
	(0,016)	(0,016)	(0,008)	(0,008)	(0,014)	(0,018)

Table 7 - Estimates for the Log(Real Daily Wages) - Equation for men born in 1958

Explanatory variables	Pooled OLS	Pooled 2SLS	Fixed Effects	FE-2SLS	Procedure B (S&W)	Kyriazidou
Unempl. duration	-0,126	-0,137	-0,072	-0,067	-0,075	-0,057
	(0,006)	(0,008)	(0,003)	(0,004)	(0,008)	(0,009)
Living in Paris	0,145	0,145	0,117	0,117	0,093	0,105
	(0,009)	(0,009)	(0,012)	(0,012)	(0,02)	(0,025)
Part-time	-0,501	-0,496	-0,345	-0,345	-0,306	-0,334
	(0,016)	(0,016)	(0,007)	(0,007)	(0,013)	(0,014)
Experience/10	0,624	0,628	0,563	0,56	0,426	0,546
	(0,026)	(0,026)	(0,013)	(0,014)	(0,022)	(0,03)
Experience²/100	-0,146	-0,147	-0,124	-0,124	-0,077	-0,121
	(0,008)	(0,008)	(0,005)	(0,005)	(0,007)	(0,01)
Manager, salaried of its own firm	0,574	0,572	0,183	0,184	0,168	0,181
	(0,038)	(0,038)	(0,019)	(0,019)	(0,031)	(0,038)
Executive	0,79	0,789	0,224	0,224	0,212	0,212
	(0,013)	(0,013)	(0,01)	(0,01)	(0,015)	(0,02)
Mid-class working	0,384	0,383	0,099	0,099	0,093	0,095
	(0,009)	(0,009)	(0,007)	(0,007)	(0,01)	(0,014)
White collar	0,092	0,092	-0,011	-0,011	-0,014	-0,008
	(0,012)	(0,012)	(0,008)	(0,008)	(0,013)	(0,016)

Table 8 - Estimates for the Log(Real Daily Wages) - Equation for men born in 1962

Explanatory variables	Pooled OLS	Pooled 2SLS	Fixed Effects	FE-2SLS	Procedure B (S&W)	Kyriazidou
Unempl. duration	-0,098	-0,11	-0,058	-0,065	-0,089	-0,044
	(0,006)	(0,008)	(0,004)	(0,005)	(0,009)	(0,011)
Living in Paris	0,131	0,131	0,118	0,117	0,099	0,106
	(0,009)	(0,009)	(0,011)	(0,011)	(0,019)	(0,024)
Part-time	-0,486	-0,482	-0,366	-0,365	-0,301	-0,346
	(0,014)	(0,014)	(0,007)	(0,007)	(0,012)	(0,014)
Experience/10	0,868	0,873	0,821	0,825	0,501	0,754
	(0,024)	(0,024)	(0,013)	(0,013)	(0,021)	(0,031)
$\text{Experience}^2/100$	-0,27	-0,273	-0,251	-0,252	-0,11	-0,23
	(0,01)	(0,01)	(0,006)	(0,006)	(0,009)	(0,013)
Manager, salaried of its own firm	0,407	0,405	0,139	0,138	0,096	0,135
	(0,036)	(0,036)	(0,022)	(0,022)	(0,031)	(0,044)
Executive	0,712	0,71	0,324	0,323	0,258	0,288
	(0,012)	(0,012)	(0,011)	(0,011)	(0,017)	(0,022)
Mid-class working	0,356	0,355	0,136	0,135	0,107	0,125
	(0,009)	(0,009)	(0,007)	(0,007)	(0,01)	(0,015)
White collar	0,082	0,082	0,013	0,013	0,006	0,012
	(0,011)	(0,011)	(0,008)	(0,008)	(0,012)	(0,016)

Table 9 - Estimates for the Log(Real Daily Wages) - Equation for women born in 1950

Explanatory variables	Pooled OLS	Pooled 2SLS	Fixed Effects	FE-2SLS	Procedure B (S&W)	Kyriazidou
Unempl. duration	-0,079	-0,085	-0,069	-0,069	-0,061	-0,061
	(0,006)	(0,007)	(0,004)	(0,004)	(0,009)	(0,011)
Living in Paris	0,184	0,184	0,205	0,205	0,168	0,213
	(0,015)	(0,015)	(0,022)	(0,022)	(0,05)	(0,048)
Part-time	-0,497	-0,496	-0,353	-0,353	-0,317	-0,328
	(0,015)	(0,015)	(0,007)	(0,007)	(0,012)	(0,015)
Experience/10	0,745	0,747	0,79	0,791	0,518	0,684
	(0,034)	(0,034)	(0,018)	(0,018)	(0,032)	(0,043)
Experience²/100	-0,121	-0,121	-0,139	-0,139	-0,078	-0,128
	(0,008)	(0,008)	(0,004)	(0,004)	(0,007)	(0,009)
Manager, salaried of its own firm	0,5	0,499	0,218	0,218	0,159	0,188
	(0,064)	(0,064)	(0,037)	(0,037)	(0,063)	(0,078)
Executive	0,776	0,775	0,123	0,123	0,119	0,126
	(0,029)	(0,029)	(0,017)	(0,017)	(0,027)	(0,035)
Mid-class working	0,436	0,435	0,054	0,054	0,049	0,06
	(0,021)	(0,021)	(0,012)	(0,012)	(0,019)	(0,025)
White collar	0,187	0,186	0,049	0,049	0,034	0,039
	(0,016)	(0,016)	(0,01)	(0,01)	(0,017)	(0,021)

Table 10 - Estimates for the Log(Real Daily Wages) - Equation for women born in 1954

Explanatory variables	Pooled OLS	Pooled 2SLS	Fixed Effects	FE-2SLS	Procedure B (S&W)	Kyriazidou
Unempl. duration	-0,084	-0,089	-0,067	-0,071	-0,069	-0,06
	(0,007)	(0,008)	(0,004)	(0,005)	(0,009)	(0,011)
Living in Paris	0,175	0,175	0,239	0,238	0,188	0,207
	(0,014)	(0,014)	(0,021)	(0,021)	(0,039)	(0,046)
Part-time	-0,533	-0,532	-0,382	-0,381	-0,338	-0,362
	(0,013)	(0,013)	(0,007)	(0,007)	(0,011)	(0,015)
Experience/10	0,723	0,725	0,827	0,83	0,525	0,715
	(0,033)	(0,033)	(0,019)	(0,019)	(0,031)	(0,045)
Experience²/100	-0,131	-0,132	-0,166	-0,167	-0,085	-0,15
	(0,009)	(0,009)	(0,005)	(0,005)	(0,008)	(0,012)
Manager, salaried of its own firm	0,627	0,626	0,189	0,189	0,17	0,203
	(0,06)	(0,06)	(0,037)	(0,037)	(0,054)	(0,079)
Executive	0,805	0,803	0,143	0,142	0,127	0,149
	(0,027)	(0,027)	(0,017)	(0,017)	(0,025)	(0,036)
Mid-class working	0,424	0,423	0,053	0,053	0,035	0,061
	(0,019)	(0,019)	(0,012)	(0,012)	(0,02)	(0,026)
White collar	0,225	0,224	0,075	0,074	0,047	0,07
	(0,016)	(0,016)	(0,01)	(0,01)	(0,018)	(0,022)

Table 11 - Estimates for the Log(Real Daily Wages) - Equation for women born in 1958

Explanatory variables	Pooled OLS	Pooled 2SLS	Fixed Effects	FE-2SLS	Procedure B (S&W)	Kyriazidou
Unempl. duration	-0,077	-0,082	-0,053	-0,055	-0,052	-0,048
	(0,005)	(0,006)	(0,003)	(0,004)	(0,007)	(0,01)
Living in Paris	0,178	0,178	0,17	0,17	0,123	0,162
	(0,012)	(0,012)	(0,019)	(0,019)	(0,03)	(0,042)
Part-time	-0,513	-0,512	-0,415	-0,415	-0,361	-0,383
	(0,012)	(0,012)	(0,007)	(0,007)	(0,011)	(0,014)
Experience/10	0,816	0,818	0,738	0,739	0,487	0,607
	(0,032)	(0,032)	(0,019)	(0,019)	(0,029)	(0,045)
Experience²/100	-0,181	-0,182	-0,168	-0,169	-0,084	-0,146
	(0,011)	(0,011)	(0,007)	(0,007)	(0,009)	(0,015)
Manager, salaried of its own firm	0,573	0,573	0,23	0,23	0,203	0,241
	(0,053)	(0,053)	(0,039)	(0,039)	(0,051)	(0,081)
Executive	0,828	0,827	0,17	0,17	0,167	0,176
	(0,023)	(0,023)	(0,018)	(0,018)	(0,027)	(0,037)
Mid-class working	0,425	0,424	0,079	0,079	0,075	0,086
	(0,017)	(0,017)	(0,012)	(0,012)	(0,019)	(0,026)
White collar	0,22	0,219	0,085	0,085	0,063	0,077
	(0,013)	(0,013)	(0,01)	(0,01)	(0,016)	(0,021)

Table 12 - Estimates for the Log(Real Daily Wages) - Equation for women born in 1962

Explanatory variables	Pooled OLS	Pooled 2SLS	Fixed Effects	FE-2SLS	Procedure B (S&W)	Kyriazidou
Unempl. duration	-0,08	-0,089	-0,049	-0,056	-0,051	-0,043
	(0,005)	(0,006)	(0,004)	(0,005)	(0,007)	(0,011)
Living in Paris	0,178	0,178	0,13	0,129	0,094	0,135
	(0,01)	(0,01)	(0,016)	(0,016)	(0,024)	(0,036)
Part-time	-0,525	-0,523	-0,43	-0,429	-0,37	-0,414
	(0,01)	(0,01)	(0,007)	(0,007)	(0,009)	(0,014)
Experience/10	0,853	0,86	0,761	0,767	0,496	0,62
	(0,029)	(0,029)	(0,018)	(0,018)	(0,026)	(0,044)
Experience²/100	-0,255	-0,259	-0,244	-0,246	-0,118	-0,208
	(0,013)	(0,013)	(0,009)	(0,009)	(0,011)	(0,019)
Manager, salaried of its own firm	0,352	0,35	0,15	0,15	0,15	0,161
	(0,055)	(0,055)	(0,042)	(0,042)	(0,048)	(0,089)
Executive	0,73	0,728	0,278	0,277	0,239	0,276
	(0,021)	(0,021)	(0,017)	(0,017)	(0,024)	(0,036)
Mid-class working	0,402	0,4	0,119	0,119	0,098	0,125
	(0,016)	(0,016)	(0,012)	(0,012)	(0,017)	(0,026)
White collar	0,233	0,231	0,109	0,108	0,08	0,102
	(0,012)	(0,012)	(0,01)	(0,01)	(0,014)	(0,021)

Table 13 – Fisher test on the Mills ratio terms for women

	Generation 50	Generation 54	Generation 58	Generation 62
Fisher statistic	90.2	112.4	135.9	169.8
Significance	<0.001	<0.001	<0.001	<0.001

Annex: descriptive statistics

	Generation 50	Generation 54	Generation 58	Generation 62
Men	7826	8056	8861	9673
Participation rate (all years)	73,6%	71,6%	73,4%	72,6%
Participation rate in 1984	75,3%	73,1%	72,4%	64,0%
Participation rate in 1992	74,8%	72,6%	74,4%	75,4%
Participation rate in 2001	69,8%	68,7%	71,2%	72,3%
Manager, salaried of its own firm	2,8%	2,1%	1,6%	1,3%
Executive	14,1%	12,9%	12,0%	10,1%
Mid-class working	23,0%	21,3%	19,6%	18,0%
Other white collar	8,7%	10,0%	10,4%	12,5%
Blue collar	51,4%	53,7%	56,4%	58,1%
Experience	19,2	15,0	11,6	8,2
Daily wage	77,4	72,1	66,9	61,4
Women	8623	8661	9372	9785
Participation rate (all years)	55,5%	56,4%	57,2%	61,0%
Participation rate in 1984	49,7%	52,1%	55,8%	60,2%
Participation rate in 1992	57,5%	57,0%	56,6%	60,0%
Participation rate in 2001	57,9%	61,9%	66,3%	68,1%
Manager, salaried of its own firm	0,8%	0,9%	0,6%	0,6%
Executive	6,2%	6,7%	6,5%	6,6%
Mid-class working	18,4%	19,3%	17,7%	16,7%
Other white collar	46,8%	48,0%	48,7%	52,3%
Blue collar	27,8%	25,1%	26,4%	23,9%
Experience	13,8	11,4	9,0	6,7
Daily wage	53,9	53,6	51,1	48,7