Sources of Wage Growth and Returns to Tenure in Italy

Michele Battisti

Department of Economics, Simon Fraser University

battisti@sfu.ca

September 24, 2009

PRELIMINARY AND INCOMPLETE, PLEASE DO NOT QUOTE

Abstract

This paper estimates the returns to experience and job tenure using a simultaneous equation model that accounts for the potential endogeneity of seniority in the wage determination process. It uses panel data for Italian workers' histories based on social security records, for the years 1985-1999. This dataset has not been used for this research purpose to this date. The hypothesis of exogeneity of seniority can be tested and is rejected due to endogeneity of the match effect. The first two years on the job are associated with a two percent yearly wage increase. Effects of seniority are very small afterwards. The paper also discusses some implications of these results.

1 Introduction

This paper uses a long panel dataset of Italian workers to estimate the return to labour market experience and seniority for Italian men and women for years 1985-1999. By doing so, the analysis also concerns the role of mobility at different levels of a person's career and the impact of mobility on wage profiles of Italian employees.

Good theoretical analysis and precise estimation of the accumulation of human capital and in particular of the returns to tenure (in the various forms in which tenure can be interpreted) is crucial for a better understanding of the prevalence of long-term matches, but also for better evaluating the welfare consequences of turnovers in the labour market, especially in a context with increasing labour market mobility. Estimating the returns to experience and seniority provides guidelines for the design of labour market programs. Assessing the costs and benefits of mobility sheds light on the mechanism underlying human capital accumulation, and helps evaluating the welfare effect of displacement for workers.

The estimation of the impact of experience and seniority on wages is carried out using a very flexible model that allows for the inclusion of random person and match effects and for endogeneity of seniority in the wage equation through the simultaneous estimation of a hazard model for employment duration. The main results show that returns to seniority are sizeable only in the first few years, and are very small afterwards.

2 Background

Understanding human capital accumulation and the sources of wage growth has arguably been one of the most debated topics in labour economics. Within the human capital literature, which flourished over the last few decades after Becker's (1964) seminal contribution, there has been much discussion concerning the distinction between general and specific human capital and their respective returns. More recently, empirical research has shown the existence of different forms of specific human capital. Researchers found evidence of the importance of industry-specific human capital over firm-specific human capital (Parent 2000). Those results are integrated with more recent evidence showing that displacements seem to be associated with a large deterioration of workers' *skill portfolio* and that industry-specific human capital seems to be of secondary importance (Poletaev and Robinson 2008).

As stressed in Dustmann and Meghir (2005), studying individual wage growth and its determinants is of crucial importance for most labour market programs. Furthermore, the focus on wage growth, returns to experience and returns to seniority offer very useful insights for understanding job mobility and its consequences in terms of benefits and costs, which is itself very useful for policy-makers.

All the main predictions of the theory of human capital rely heavily on the definition of general and specific human capital. Therefore, investigating the returns to experience and to tenure (job tenure, firm tenure, occupation tenure, industry tenure, sector tenure) provides crucial insights about the empirical importance and role of different types of human capital.

The interest in the characteristics and evolution of specific human capital has resulted in a very large body of empirical research attempting to estimate returns to experience, returns to seniority and the impacts of mobility. The debate focusing on empirical issues arising with the estimation of returns to experience and seniority has been a long and fruitful one. Seminal earlier contributions from the late 1980s and early 1990s include: Altonji and Shakotko (1987), Brown (1989), Topel (1991) and Neal (1995). These contributions all recognize that an Ordinary Least Squares (OLS) regression of wages on tenure suffers an endogeneity problem whenever any unobservable characteristic of the firm, job and/or worker affects both wage and job duration. A range of solutions has been offered, each with some limitations.

After these initial contributions, and largely thanks to advancement in the econometric theory and techniques involved in the estimations, as well as to the availability of better data for addressing these issues, more reliable estimates have been produced. Furthermore, the set of questions addressed is now larger, especially due to the increased availability of integrated employer-employee data. The estimation of the return to tenure can now be carried out in models where there are worker, firm and match fixed effects, in which returns to human capital and experience may be heterogeneous, and where exogenous demand effects can be taken into account.

The complex relationship between tenure, productivity and wages is investigated by Dohmen (2004) using Dutch data, finding no evidence for strong impact of performance controls on seniority-wage profiles. More attention has also been paid to the assumptions made about observability of workers', firms' and match's characteristics by agents *vis-a-vis* the econometricians (see Gibbons, Katz, Lemieux, and Parent 2005). Also, matched employer-employee data made it possible for researchers to allow more detailed analysis of the implications of the human capital theory and of labour search and match theories (Jovanovic 1979). A very interesting example in this sense is Nagypal (2007), which tries to distinguish between two ways in which one can interpret the endogeneity between tenure and the hazard rate of separation, i.e. specific human capital accumulation on the one hand, and increased information about match quality on the other. This is also a good example of the tension between a pure interpretation of human capital theories and of matching theories.

This paper will be in some sense agnostic about the mechanisms that underlie the relationship between tenure and wage growth, and focuses on empirics. The primary goal is to estimate the importance of returns to tenure[1] in the Italian labour market, and further goals include isolating the relative importance of mobility across occupations (within the same employer), firms (changing firm, not occupation), sector and geographic area. Movements to and from unemployment can also be taken into account.

Given this focus, the most closely related literature contributions are Dostie (2005) and

[1]Tenure and seniority will be used interchangeably here

Dustmann and Meghir (2005). Dostie (2005) uses French data and estimates returns to seniority in a model where he can allow for correlated individual and job unobserved heterogeneity. He finds no significant returns to seniority after controlling for matching considerations. Dustmann and Meghir (2005) focus on similar issues but frame their analysis much more as a study of wage impacts of different sources of human capital[2]. They use German data and distinguish between skilled and unskilled workers, and find that returns to firm tenure are positive for all workers, and larger for unskilled workers. Returns to sector tenure are positive for skilled workers, but are not significantly different from zero for unskilled workers.

This work will attempt at addressing these issues using Italian data. The focus here will be on firm and sector tenure, on the wage impacts of mobility (not only between jobs and sector, but also geographic) and on the impacts of mobility on the life-cycle earnings of workers. The next section briefly outlines the empirical methodology that will be implemented in the paper.

3 Endogeneity of Job Seniority

Let $i = \{1, ..., N\}$ be a worker, and $t = \{1, ..., T\}$ be a time period. $J(i, t)$ is then the employer of worker i at time t[3].

Using a linear framework, many studies estimated an equation of the form:

$$ln(w_{iJ(i,t)t}) = X_{it}\beta_0 + \beta_1(seniority_{iJ(i,t)t}) + \beta_2(experience_{it}) + \epsilon_{iJ(i,t)t} \quad (1)$$

where

[2]The difficulties in measuring labour productivity makes it hard to distinguish between human capital accumulation and informational aspects of matches empirically

[3]In the empirical section, $J(i, t) \equiv j$ for simplicity, given the structure of the data used here, which does not allow the identification of firms

- $w_{iJ(i,t)t}$ is the real wage of worker i working for firm $J(i,t)$ in period t;
- $seniority_{iJ(i,t)t}$ is equal to the duration of the match $J(i,t)$ up to period t;
- $experience_{it}$ is equal to the total labour market experience of worker i.

The error term $\epsilon_{iJ(i,t)t}$ can be decomposed into a person fixed effect component, a firm fixed effect component, a match component and a component that will be match-, time- and person-specific:

$$\epsilon_{iJ(i,t)t} = \theta_i + \psi_{J(i,t)t} + \delta_{iJ(i,t)} + \nu_{iJ(i,t)t} \tag{2}$$

OLS yields unbiased estimates of β_0 and β_1 in (1) only if both experience and seniority are uncorrelated with unobserved heterogeneity at the individual, firm or match level (see Dostie 2005), which have an impact on the dependent variable.

Many studies have presented theoretical and empirical reasons why person, firm and match effects might be important in reality. The job search literature provided a theoretical framework in which to critically think of the relationship between wages, tenure and mobility decisions. Continuing matches are a self-selected subset of previously existing matches. For example, more productive firms might attract more productive or more motivated workers; matches might survive only if workers are productive; employers might use wages as a level for retaining employees (Dostie 2005). In all these cases, job seniority is correlated with individual, firm and match characteristics. The endogeneity of seniority cannot easily be solved by instrumenting for it (as commonly pursued in earlier work), given that the source of endogeneity is likely to come from unobservable characteristics of either worker, firm of match.

3.1 Wage equation and employment duration

The recent literature, starting with Lillard (1999), has attempted to exploit the advantages of longitudinal datasets to estimate this endogeneity problem and control for endogeneity directly. A person effect and a match effect can be estimated by modelling wages and employment duration simultaneously, as in Dostie (2005). Using panel data on work histories as in this paper does not allow the econometrician to include firm fixed effects, but a similar empirical strategy as in Lillard (1999) can be adopted nevertheless.

The match effect will include heterogeneity coming from the firm size, and thus not including a firm component impacts the interpretation of the match effect but does not affect the remaining estimates. Dostie (2005) applies a very similar methodology: even if the data used include linked employer-employee information, the author only includes a match effect.

The simultaneous model estimated in the following sections includes a wage equation and a model for employment duration. Wages are only observed for accepted job offers, and therefore represent the outcome of a sample selection process by which employments spells terminate if the worker does not accept the wage offer. Quits and layoffs are regarded as equivalent in this context, although it is necessary to acknowledge that there may be very different motivations for the termination of a job. The theoretical starting point for this analysis is a simple job search model in which on-the-job search is possible, as in Pissarides (1994). The model for employment decision is presented in the following, simplifying the analysis in Lillard (1999).

3.2 Modeling employment duration

A worker i works at firm j at time $t-1$. At the beginning of period t, she is offered a wage w_{ij_0t}, which she will accept and continue working for firm j_0 if the offered is higher than some reservation wage w^R_{ijt}. This reservation wage may be a function of personal characteristics, experience, and tenure in firm j_0. The crucial point here is that we observe w_{ij_0t} only if it is such that both parties found it optimal to continue the relationship.

Suppose for example that a worker gets an offer from a different firm j_1. This offer will have an impact on w^R_{ijt}. Let this offer be w_{ij_1t}, and assume it depends on the individual specific component (which depends on unobservables), on experience and on the match component in the following way:

$$ln(w_{ij_1t}) = f(experience_{it}) + \theta_i + \delta_{ij_1} \tag{3}$$

Also, assume the reservation wage for worker i at firm j_0 is

$$ln(w^R_{ijt}) = ln(w_{ij_0t}) + g(experience_{it}, seniority_{ijt}) \tag{4}$$

Person i will decide to move if

$$ln(w_{ij_1t}) > ln(w^R_{ijt}) \tag{5}$$

Assume she will stay in case of indifference. Substituting, i will move if:

$$\begin{aligned} f(experience_{it}) + \theta_i + \delta_{ij_1} &> ln(w_{ijt}) + g(experience_{it}, seniority_{ijt}) && (6)\\ \implies \delta_{ij_1} &> \Delta && (7) \end{aligned}$$

where

$$\Delta \equiv ln(w_{ijt}) + g(experience_{it}, seniority_{ijt}) - f(experience_{it}) - \theta_i$$

Let p_t be the probability of person i receiving on offer at all at the beginning of period t. This probability is taken as exogenous here for simplicity. Then the conditional probability of observing a job separation can be written as:

$$h(ln(w_{ijt}), experience_{it}, seniority_{ijt}) = p_t[1 - F(\Delta)] \tag{8}$$

where F is the cumulative distribution function of δ_{ij_1}. This assumes that p_t and δ_{ij_1} are uncorrelated. Intuitively, equation (8) means that the probability of a match destruction in this framework is given by the probability of having an offer at all times the probability of that offer being large enough to be accepted. This model can easily accommodate the inclusion of unemployment as an outside option that leads to workers rejecting very low offers.

4 Empirical model

4.1 Derivation

4.1.1 Hazard model

Employment duration is estimated following the discussion above, and using a hazard model based on Kiefer (1988). The hazard duration dependence follows a generalised Gompertz distribution, i.e. a piecewise linear spline. Pollard and Valkovics (1992) and Lillard and Panis (2003b) provide additional information about the characteristics and common notational conventions for this distribution.

For person i, I use a proportional hazard model of the following form:

$$\begin{aligned} ln(h_{ij}(\tau)) = \quad & \beta_0 + \beta_1(male_i) + \beta_2'\mathbf{jobcov_{jt}} + \beta_3'\mathbf{seniority_{ijt}} + \beta_4'\mathbf{experience_{ijt}} \\ & + \beta_6 time_t + \theta_{1i} + \phi\delta_{ijt} \end{aligned} \tag{9}$$

where

- $male_i$ is the only personal characteristics considered here[4];
- $\mathbf{jobcov_{jt}}$ is a vector of job characteristics, such as occupation, sector, region etc.;
- $time_t$ is a linear time trend;
- θ_{1i} is a random person effect;
- δ_{ijt} is a random match effect, with load parameter ϕ
- all random effects have zero mean;
- $\mathbf{seniority_{ijt}}$ and $\mathbf{experience_{ijt}}$ are piecewise linear splines, so that a spline $\Upsilon(\tau)$ with $\eta_{\mathbf{N}}$ nodes is based on the following transformation of the spell duration:

$$\Upsilon(\tau) = \begin{cases} min[\tau, \eta_1] \\ max[0, min(\tau - \eta_1, \eta_2 - \eta_1)] \\ ... \\ max[0, min(\tau - \eta_{N-1}, \eta_N - \eta_{N-1})] \\ max[0, \tau - \eta_N] \end{cases}$$

The survivorship function for equation (9) without time-varying covariates[5] is given by the following equation (also see Lillard and Panis (2003b)):

$$S_i(\tau) = exp\left(-\int_0^\tau h(\iota)d\iota\right) = exp\left(-\int_0^\tau e^{\alpha' T(\iota)}d\iota\right)^{e^{\beta' male}} \tag{10}$$

where T includes the piecewise linear splines of equation (9).

[4]Education is not available, and age seem to be inappropriate in this model given the sample selection that will be made

[5]As it is the case for the equation estimated in this version of the paper

4.1.2 Wage equation

The wage equation is specified as in the following:

$$
\begin{aligned}
ln(w_{ijt}) = \quad & \gamma_0 + \gamma_1(male_i) + \gamma_2'\mathbf{jobcov_{jt}} + \gamma_3'\mathbf{seniority_{ijt}} \\
& +(1 + \gamma_4 male_i + \theta_{3i})\gamma_5'\mathbf{experience_{ijt}} + \gamma_6 time_t \\
& +\theta_{2i} + \delta_{ijt} + \nu_{ijt}
\end{aligned} \tag{11}
$$

where, as before, the subscript ijt is used for simplicity instead of $iJ(i,t)t$, which would have been redundant notation given that I never specify firm effects and match effects separately. In equation (11), θ_{2i} and θ_{3i} are random person effects and ν_{ijt} is the person-match-time specific error term. All other covariates are defined as in equation (9). All random effects have zero mean. Equation (11) includes random person and match effects, and also allows personal characteristics and unobserved heterogeneity at the person level to impact the returns to experience.

4.1.3 Assumptions on parameters

I assume a first-order autoregressive form for the error term in equation (11):

$$
\nu_{it} = \omega \cdot \nu_{i,t-1} + u_{it} \tag{12}
$$

where $u_{it} \sim iid\ N(0, \sigma_\delta^2)$. This allows for person-specific correlation among all wage values within a worker's career, which holds across jobs.

The model also allows for job-specific heterogeneity in the initial wage level, which is independent from job to job. For every match j

$$\delta_{ijt} \sim iid\ N(0, \sigma_\delta^2)$$

4.1.4 Simultaneity

Two sets of elements introduce simultaneity in the two-equation model described above. Firstly, the individual random effects components of equation (9) and of equation (11) are jointly normally distributed, i.e.

$$(\theta_{1i}, \theta_{2i}, \theta_{3i}) \sim N(\mathbf{0}, \mathbf{\Sigma}_{\theta,\theta})$$

Secondly, a load factor ϕ on the match random effect component of equation(11) is introduced in equation (9).

This model allows us to test exogeneity of seniority in the wage equation, and to test some of the predictions of a matching model of the labour market. Two tests for exogeneity are available here, as either coming from the person unobserved heterogeneity or from the match unobserved heterogeneity. The null hypothesis of zero correlation between θ_{1i} and θ_{2i} or θ_{3i} will be rejected in case of endogeneity. Exogeneity of seniority in the wage equation would also imply $\phi = 0$, which is the second way of testing exogeneity.

4.2 Estimation

It is possible to derive the full joint marginal likelihood function for the joint model of wage determination and job duration. For this it is crucial that person and match effects are nested, which allows me to adapt the analysis in Lillard (1999) to the hazard and the wage equations estimated in this paper. It is necessary to assume independence of individual likelihoods conditional on person and job heterogeneity (Dostie 2005).

Under the assumption stated above, the full joint marginal likelihood for person i can be expressed as:

$$
\begin{aligned}
L_i = \quad & (2\pi)^{\frac{-T_i}{2}} |\mathbf{\Sigma}_{\zeta_\mathbf{i},\zeta_\mathbf{i}}| e^{\left[-\frac{1}{2}(\mathbf{W}_i - \mathbf{\Xi}_i)' \mathbf{\Sigma}^{-\mathbf{1}}_{\zeta_\mathbf{i},\zeta_\mathbf{i}} (\mathbf{W}_i - \mathbf{\Xi}_i)\right]} \\
& \times \int_{-\infty}^{\infty} (2\pi\sigma^2_{\theta_1|\mathbf{W}_i})^{-\frac{1}{2}} e^{-\frac{1}{2}\left(\frac{\theta_1 - \mu_{\theta_1|\mathbf{W}_i}}{\sigma_{\theta_1|\mathbf{W}_i}}\right)^2} \\
& \times \left\{ \int_{\delta_{iJ}} (2\pi\sigma^2_{\delta_{iJ}})^{-\frac{1}{2}} e^{-\frac{1}{2}\left(\frac{\delta_{iJ} - \mu_{\delta_{iJ}|\mathbf{W}_i}}{\sigma_{\delta_{iJ}|\mathbf{W}_i}}\right)^2} [(h_{\theta_{1i},\delta_{iJ}}(\iota_{iJ}))^{D_J} S_{\theta_{1i},\delta_{iJ}}(\iota_{iJ})] d\delta_{iJ} \right. \\
& \left. \times \prod_{J=1}^{J_i - 1} \int_{\delta_{iJ}} (2\pi\sigma^2_{\delta_{iJ}})^{-\frac{1}{2}} e^{-\frac{1}{2}\left(\frac{\delta_{iJ} - \mu_{\delta_{iJ}|\mathbf{W}_i}}{\sigma_{\delta_{iJ}|\mathbf{W}_i}}\right)^2} [(h_{\theta_{1i},\delta_{iJ}}(\iota_{iJ})) S_{\theta_{1i},\delta_{iJ}}(\iota_{iJ})] d\delta_{iJ} \right\} d\theta_1
\end{aligned} \tag{13}
$$

where D_j is equal to one if the last job of person i has ended, and is equal to zero if it has not.

In the following, the components if equation (13) are defined and explained. Let $\mathbf{W}_i$ be the vector of wage values observed for person i. $\mathbf{W}_i$ can be divided into mean sub-vectors and residual sub-vectors:

$$\mathbf{W}_i = \mathbf{\Xi}_i + \zeta_i$$

Mean sub-vectors for job J are given by the following regression equation:

$$
\begin{aligned}
\Xi_{iJ(i,t)} = \quad & [\mathbf{1}_{T_{iJ(i,t)}}, \gamma_5' \mathbf{experience}_{it}, \gamma_6 time_t] \begin{pmatrix} \gamma_0 & \gamma_1 \\ 1 & \gamma_4 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ male_i \end{pmatrix} \\
& + [\mathbf{1}_{T_{iJ(i,t)}}, \gamma_3' \mathbf{seniority}_{iJ(i,t)t}] \begin{pmatrix} 0 & \gamma_\mathbf{2}' \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{jobcov_{J(i,t)t}} \end{pmatrix}
\end{aligned}
$$

Residual sub-vectors (related to the person and match random effects and to the error term) are given by:

$$\zeta_{ij} = [\mathbf{1}_{T_{iJ(i,t)}}, \gamma_5'\mathbf{experience}_{it}] \begin{pmatrix} \theta_{2i} & \delta_{iJ(i,t)} \\ \theta_{3i} & 0 \end{pmatrix} + \nu_{iJ(i,t)}$$

Consequently, the covariance matrix of the residuals for all T_i wage observations of person i is given by:

$$\mathbf{\Sigma}_{\zeta_\mathbf{i},\zeta_\mathbf{i}} = [\mathbf{1}_{T_i}, \gamma_5'\mathbf{experience}_{it}]\Sigma_{\theta,\theta}[\mathbf{1}_{T_i}, \gamma_5'\mathbf{experience}_{it}]' \; + \sigma_\delta^2 \cdot I_{J_i \times T_i} + diag_{J_i}\mathbf{\Sigma}_{\nu_\mathbf{i},\nu_\mathbf{i}}$$

with ${}_nu$ following equation (12).

Conditional on $\mathbf{W_i}$, the mean of the person-specific component entering the job duration equation, is:

$$\mu_{\theta_{1i}|\mathbf{W}_i} = \mathbf{\Sigma}_{\theta_{\mathbf{1i}},\zeta_\mathbf{i}}\mathbf{\Sigma}^{-\mathbf{1}}_{\zeta_\mathbf{i},\zeta_\mathbf{i}}(\mathbf{W}_i - \mathbf{\Xi}_i)$$

Its variance is given by:

$$\sigma^2_{\theta_{1i}|\mathbf{W}_i} = \sigma^2_{\theta_{1i}} - \mathbf{\Sigma}_{\theta_{\mathbf{1i}},\zeta_\mathbf{i}}\mathbf{\Sigma}^{-\mathbf{1}}_{\zeta_\mathbf{i},\zeta_\mathbf{i}}\mathbf{\Sigma}_{\zeta_\mathbf{i},\theta_{\mathbf{1i}}}$$

where

$$\mathbf{\Sigma}_{\zeta_\mathbf{i},\theta_{\mathbf{1i}}} = [\mathbf{1}_{T_i}, \gamma_5'\mathbf{experience}_{it}] \begin{pmatrix} \sigma_{\theta_{1i},\theta_{2i}} \\ \sigma_{\theta_{1i},\theta_{3i}} \end{pmatrix}$$

Estimation of this model cannot be carried out using any of the most common commercial statistical software packages. However, aML (Applied Maximum Likelihood), a software developed by late Lee A. Lillard and Constantijn W.A. Panis (Lillard and Panis 2003a), can estimate multilevel nonlinear models of this kind.

The multilevel correlated random effect model is estimated under the assumption that the heterogeneity components and the innovation component in the error's AR(1) process are jointly normally distributed. The estimation is based on the maximisation of the marginal likelihood.

5 The data

5.1 The WHIP Dataset

The empirical work carried out for this paper is based on the *Work Histories Italian Panel* (WHIP in the following)[6]. WHIP is a database of individual work histories, based on administrative archives from the *Istituto Nazionale della Previdenza Sociale*[7] (INPS in the following), which is the main institution for social security in Italy. By the Italian labour market law, it is compulsory to obtain insurance from INPS for all employees in the private sector, for some categories of employees of the public sector and for most self employed[8]. Detailed descriptions of the WHIP dataset are available from Contini (2002) and Contini and Trivellato (2005).

The reference population of WHIP is made up by all the people Italian-born and foreign-born who have worked in Italy in any of the years of the panel. From this population, a sample has been extracted at random, using four dates of birth for workers for each year. This results in a dynamic population of about 370,000 people in the freely accessible version of the dataset and around twice as many in the full version[9].

For all individuals, the main episodes of their working careers are observed. The complete

[6]WHIP Work Histories Italian Panel, is a database of work histories developed by Laboratorio Revelli Centre for Employment Studies. See http://www.laboratoriorevelli.it/whip

[7]National Institute for Social Security

[8]The only notable exceptions being professionals, such as doctors, lawyers and notaries

[9]This version of the paper uses the free release of the dataset

list of observations includes: private employee working contracts, atypical contracts, self-employment activities as artisan, trader and some activities as freelancer, retirement spells, as well as non-working spells in which the individual received social benefits, e.g. unemployment subsidies or mobility benefits. The workers for whom activity is not observed in WHIP are those who do not receive insurance from INPS (and those for which it was not possible to construct a representative sample from INPS data): most public sector employees and freelancers such as lawyers, notaries and journalists, who have alternative social security funds. The black market, which represents a significant share of the Italian labour market, is also absent.

Individual data include gender, year and region of birth, received unemployment compensations or enrolment in other programs[10]. WHIP does not include educational attainments.

This paper will use the section of WHIP that concerns employees, for which the database also provides some information about employers, such as firm size, location (five regions), industrial sector (18 sectors in the standard version of WHIP, 34 in the full version)[11]. Information about firms are rare in datasets of this nature, and represent a very promising source for addressing a number of questions of interest to researchers and policy-makers.

WHIP includes a series of information about *jobs*[12], such as match duration (dates of match creation and match destruction), wage received[13], other benefits received by the employee, special arrangements, occupation, location (which may be different from that of the employer). The observed period goes from 1985 to 1999, and will be regularly upgraded by the researchers of *Laboratorio Revelli*.

[10] In this version, only gender is included

[11] See Contini (2002) for further details on the criteria used for identifying firms from an economist perspective in WHIP. The researchers of Laboratorio Revelli that created WHIP did a very extensive work with the identification of firms in an economic sense, and not in a legal sense that is identified by INPS data. This makes sure that M&As, changes of names or legal headquarter do not impact the analysis

[12] *Job* (used interchangeably with *match*) is intended here as a specific employer-employee relationship, and thus to be clearly distinguished from *occupation*

[13] Net of all employer's contributions

5.2 Sample selection

The full sample of the WHIP dataset is not appropriate for being used for the estimation of the returns of experience and seniority. The most important reason is that there is no information regarding human capital accumulation of individuals (specifically, schooling and training) nor on labour market experience before 1985. Therefore, using the full dataset would only be possible by constructing of a variable measuring *potential experience* as function of age. This paper attempts instead to select the appropriate sample to avoid the construction of this variable and its arbitrariness.

Only workers for whom it is sensible to assume that we can observe them from the beginning of their careers are included. For this purpose, I eliminate all workers that are matched in 1985, the first year of the WHIP panel. I also eliminate people that are over 25 years of age in 1986. My sample is therefore represented by people that are born in 1961 and later, and that were not working in 1985. I will assume that the first time these workers are observed in the sample (in 1986 or later), they are starting their careers. The interpretation should take account of the fact that all results are based on a population of workers that are on average much younger than the overall Italian labour force. The oldest worker I have in my sample is 25 years of age is 1986, and thus 38 in 1999.

Other modifications to the data have been carried out before constructing summary statistics and estimating the model described above. Only nominal wages are included in the WHIP dataset, and so I convert all wage measures into year-1999 Euros, by using aggregate CPI data from ISTAT[14]. Given that WHIP includes number of *Full Time Equivalent* (FTE) days for each worker in each year, I constructed annual FTE wages, which have much higher comparability across workers than total yearly wages directly available from WHIP.

A second issue for the application of the WHIP dataset to this research question comes from the different length of jobs. WHIP includes information about start date and end date

[14] *Istituto Nazionale di Statistica*, the Italian National Statistical Agency

of each job. These may or may not go across years. However, in order to avoid overweighting observations for short employment spells (there can be more than one of those for each worker in each year) and in order to avoid imputations based on yearly observations of wages (only total wages for each year in each job are included), in this version of the paper I use yearly data. For every worker and every year, a *dominant* job is identified. First, I eliminate all jobs with less than five *full-time equivalent* working days. Then I rank jobs by number of effective full-time-equivalent days, and then by duration and wages. Only the first job of this ranking is kept for every worker in every year.

The dataset used for the regressions, which does not include firm size, region, occupation and sector, and thus does not take into account missing values in those variables, consists of 48,475 workers, 101,917 job spells, 252,586 yearly wage observations.

5.3 Summary Statistics

This section is divided into summary statistics information pertaining to the worker, job and annual levels.

5.3.1 Workers

This section provides summary statistics concerning the sample of workers used in this paper. Given that the goal here is to illustrate the characteristics of the sample, rather than providing an overview of the Italian labour market, all statistics in this section are based upon the same dataset that is used in the regressions. The number of observations is slightly lower for some of the tables. This is simply due to some missing values. Data is this subsection are based on workers' first job only, when this is relevant. This is done for clarity, and these distributions are not sensitive to this choice.

In the sample used here, 61 percent of the workers are male, 39 percent are female. They are much younger than the average worker in the Italian labour market. As a result of the restrictions imposed here, most of the workers in the sample are observed in a match for the first time between the age of 20 and 25, as shown in Figure 1.

Table 1 shows the distribution of workers of the dataset used for regressions by region of birth. It shows that immigrants are over-represented in comparison to the overall Italian labour market of the same period. The primary cause for this is likely to be the age structure of the sample. Combining Table 1 and 2 illustrates in a summary fashion the migration pattern of workers born in the South and on the two large Italian islands (Sicily and Sardinia) to the North of the country[15].

The WHIP dataset that is used here differentiates between five occupational categories of workers. The distribution of workers by occupation is shown in Table 3, with data based on each worker's first job in our sample. Virtually all of the workers in our sample are in lower-level occupations, i.e. apprentices, workers or clerks. This is due partly to the employment insurance system in Italy, and partly to the sample selection process including only workers at the start of their careers. Around 90 percent of the workers are in a full-time job and 10-percent workers are in a part-time job.

Table 4 shows the distribution of workers by sector of the firm for which they work. *Manufacturing* is the largest sector (39 percent of workers), followed by *Wholesale and Retail Trade* (17 percent) and by *Construction* (12 percent). Table 5 breaks down that distribution along gender. Females are, relative to males, more concentrated in the *Wholesale and Retail Trade* sector, in the *Hotel and Restaurant* sector, and in other services (*Computing, rental, research and other business* and *Other Social and personal services* sectors).

Given that the empirical strategy used here relies on identifying person and job random effects through mobility, it is crucial to provide some information on how much mobility is observed. Graph 2 shows that, taking account of the length of the panel, there is a relatively

[15] Much more detailed information about mobility could be extracted from this dataset. It is not the focus of this paper

large degree of mobility in the data, as stressed by Contini and Trivellato (2005) for the Italian market as a whole. We observe one employment spell for 47 percent of the sample, two spells for 25 percent of the sample, three spells for 14 percent of the sample. Around 15 percent of the workers in our sample have four jobs or more in the 14-year period considered here[16]. These high mobility figures are largely due to the fact that workers in this sample are in the first years of their careers.

5.3.2 Jobs

The unit of observation for the statistics presented in this section is the employment spell[17].

Table 6 shows that in 63 percent of employment spells the worker is a male. Employment spells last on average just under two years. In the sample used here, 23 percent of the spells are censored, i.e. the employment relationship is continues until the last period of our panel. Workers enter employment spells with 1.11 years of experience on average. Table 7 and Table 8 present separate information for female and for male workers. In this sample, females stay on the job slightly longer than males and enter an employment spell with slightly less experience, although these differences are not statistically significant.

5.3.3 Annual

In the sample used here, we observe wages every year. The summary statistics below are based upon real wages in 1999 Euros. As shown in Table 9, workers in this sample receive on average around 16,500 Euros in yearly wages, calculated on a full-time full-year equivalent, so that spells of shorter lengths and part-time jobs are comparable. Real wages are top-coded at around 100,000 Euros[18]. At the start of the year, workers have on average 2.68 years of experience, and have accumulated tenure on the job of 1.64 years.

[16]It is important to stress that very short employment spells and of course jobs that are not recorded by INPS are excluded

[17]"Employment spell", "Job" and "Match" are used interchangeably here

[18]Total annual wages are also top-coded, and thus so are the FTE wages constructed here

5.3.4 Earnings Profiles

Figure 3 shows that there is a strong positive unconditional correlation between experience and wages. Early studies using OLS regressions were picking up this correlation and interpreting it as a causal relationship. It is interesting to note the decline for the highest level of experience observed. The most likely interpretation of this, which is visible in the equivalent graph for seniority earning profile (Figure 4), is that the set of jobs for which experience and seniority are highest (i.e. jobs that existed for the whole period studies here) are special and different from the others, possibly due to composition by region and sector, and on average pay less than matches with lower levels of experience and tenure.

Finally, an unconditional time-earning profile is presented in Figure 5. This is constructed from average real wages for females and males in each year. It clearly includes a positive time trend as well as business cycle fluctuations. Wages seem to vary over time in a very similar manner for women and men, suggesting that changes in real wages through changes in inflation may be sizeable.

6 Results

Regression results are presented in three separate tables. Table 10 presents the results for the hazard model of employment duration; Table 11 presents the estimates for the wage equation; Table 12 presents the estimates for the variance-covariance matrix of the heterogeneity components and of the error structure. Column *SM* in all three tables presents the estimates for hazard model and the wage equation estimated simultaneously.

6.1 Employment duration

Table 10 presents all the results for the hazard equation. Model H1 does not include individual heterogeneity, which is introduced in H2. Column SM refers to the estimates of the simultaneous model. The choice of nodes for the piece-wise linear splines for experience and seniority are based on best fitting the data[19].

For males, the conditional probability of match destruction in any year is around 10 percentage points higher than for females, and this estimate is robust to including individual heterogeneity and/or estimating the model simultaneously. *Ceteris paribus*, probability of job destruction increases by around two percent a year between 1986 and 1999, as shown by the coefficients on the time trend component. The impacts of seniority and experience are as expected. The first two years of seniority are associated with a lower probability of match destruction. However, the estimates are much lower once individual heterogeneity is introduced, falling from 42 percent to 13 percent. After the first two years seniority has a much smaller impact on the probability of match destruction. The effect of seniority on employment duration is also highly nonlinear. Very similar considerations arise from looking at the effects of experience on the probability of match destruction, with a very large effect for the first years and a lower effect for higher levels. All these results are similar to the results in Dostie (2005) who uses French data.

6.2 Wage Equation

Table 11 is organised as follows: column W1 presents estimates of equation (11) not including any unobserved heterogeneity component; W2 introduces an intercept and slope random effect component for individual heterogeneity; W3 also includes a match heterogeneity component; SM, as above, presents estimates for the simultaneous model accounting for endogeneity of seniority and experience in the wage determination process.

Males earn consistently higher wages than females in this sample: their conditional wages are almost five percentage points higher. Interestingly, this result is robust to the introduction of person and match fixed effects suggesting that the source of the effect is not personal

[19] Lillard and Panis (2003b) discusses the choice of nodes in further details

unobserved heterogeneity. This is in contrast with Dostie (2005) where males have lower baseline wages than females. Table 11 shows that males also have significantly higher returns to experience. For example, looking at column SM the yearly return for the first four years of experience is 4.5 percent for females, and 5.5 percent for males[20]. The time trend shows that real wage growth has been positive but very small in this period, estimated as 0.53 percent a year in the simultaneous model.

Focusing on column SM, it is clear that after the first four years experience does not have high returns in terms of wages. The average return to one year of experience within the 4-8 band is 0.88 percent for females, and around one percent for males. After the 8th year, returns are a little higher, at 1.4 percent for females, 1.7 percent for males. As in Dostie (2005) the estimates of the returns to experience are not strongly affected by enriching our specification.

The returns to seniority estimated in column W1 are around 2.9 percent a year for the first two years, not significantly different from zero for years 2-4 and 8+, and are positive for years 4-8. These estimates decrease once we introduce heterogeneity at the person and match level (especially for the returns to experience after the first two years), but by far less than in Dostie (2005), where they become negative for all years in the simultaneous regression. My estimates of the return to seniority in the simultaneous model are 2.1 percent for the first two years, negative for the following two years and very close to zero afterwards. These estimates are more similar to those of Abowd, Kramarz, and Margolis (1999).

6.3 Heterogeneity components and error structure

Given the involved structure of the model, the estimates concerning the variance-covariance matrix of the heterogeneity components for person and match effects are presented separately in Table 12.

[20]Calculated as $(1 + 0.2124) \times 0.0450$

The second column presents results from models H1 and W2 estimated separately, and thus only includes individual-level heterogeneity components in the wage equation. Both intercept and slope individual unobserved heterogeneity are found to be significant at the one-percent level. σ_{θ_3} is also very large, showing the importance of allowing for wage growth to be different across individuals as a function of unobservables.

The third column of Table 12 adds an individual heterogeneity component to the hazard model and a match heterogeneity component to the wage equation. They are both large and highly significant, and have an impact on the other coefficients as discussed above. In particular, the seemingly highly autoregressive nature of the error term of the wage equation drops dramatically as a consequence, as shown by the fall of ω from 0.66 in column H1+W1 to 0.11 in column H2+W3.

Finally, the comparison between the third column and column SM, which refers to the simultaneous model, allows me to test endogeneity of seniority in the wage equation. As discussed above, two tests are available here. Using the correlation between the person effect in the hazard equation and in the wage equation I fail to reject the null of exogeneity. This result is in contrast with results in Dostie (2005), Abowd, Kramarz, and Margolis (1999) and Lillard (1999). There is no evidence of endogeneity through the correlation of the person random effects in the hazard model and in the wage equation.

On the other hand, the estimate for ϕ in the hazard model, the coefficient of the match random effect from the wage equation, is negative and statistically significant at any conventional significance level. The estimated coefficient implies that a job that has a match effect one standard deviation higher than zero in the wage equation, has a predicted probability of destruction 14.34% lower[21] than a job with the average match effect. "Good" job matches, i.e. matches with ceteris paribus high starting wages, and/or matches in "good" firms also last significantly longer[22]. Longer-lasting matches are not a random sample of all matches

[21] $(-0.7739) \times (0.1954)$

[22] Given that firm effects are not included here, this match effect would include both a firm effect and a

in terms of wages; they tend to have on average much higher wages. On the basis of this result, we can reject the null hypothesis of exogeneity of seniority in the wage equation. An empirical model that ignores the endogeneity of job duration will yield biased estimates of the return to seniority.

7 Concluding remarks

There are solid theoretical reasons to believe that seniority and wages are simultaneously determined. The searching and matching literature offers a theoretical framework for understanding wage determinations and employment duration that takes this simultaneity into account. This paper uses data for a sample of Italian workers in years 1986-1999 to estimate the effect of seniority on wages taking account of the endogeneity of seniority in the wage equation. To the best of my knowledge, this is the first estimation of the returns to seniority using Italian panel data.

The simultaneous estimation of the wage equation and of a hazard model for employment duration produces estimates that are not biased by endogeneity problems. Through this estimation, some of the predictions of the matching literature can be evaluated. The methodology employed here also allows the inclusion of experience and seniority as piecewise linear splines, so that their effect can very over time and provide us information on nonlinearity in the response.

Results show that in the Italian labour market, which is considered among the most rigid in the OECD countries (Contini and Trivellato 2005), job search and job match considerations are important determinants of wages. Wage determination does not seem to be entirely explained by collective bargaining, as stressed also by Contini, Leombruni, Pacelli, and Villosio (2007). Individual unobservables have a large effect both on the probability of

pure match effect

job destruction, on the wage levels and on its growth. The match heterogeneity component is significant in the wage equation. The negative correlation of the match random effect component between the wage equation and the hazard equation is clear evidence of the endogeneity of seniority in the wage equation coming through unobserved match heterogeneity. "Good" jobs pay more and last longer.

The simultaneous model shows that for this sample there are short-term returns to seniority, at around two percent a year for the first two years, but after the first two years returns are very small or negative. The effects of experience on wages are also concentrated in the first years in the labour market, and are very small afterwards. The implied losses for young workers from match destruction appear to be moderate. Also, programs trying to create stable matches by subsidising firms hiring young workers seem to have low returns in the terms of future wages of those workers.

The WHIP dataset offers possibilities for expanding this research in a number of directions. A future release of the data will include more recent years, more detailed information on the characteristics of the firm for which each worker works, and a firm identifier will also be added. Furthermore, information about sector and size of the firm has the potential of deepening our understanding of the roots of returns to tenure, in particular distinguishing between job-specific productivity and reasons relating to informational problems. Finally, it would be of great interest to look beyond the average returns to seniority, investigating the *distribution* of the returns to seniority across sectors, occupations, and regions.

APPENDICES

A Summary statistics

A.1 Workers

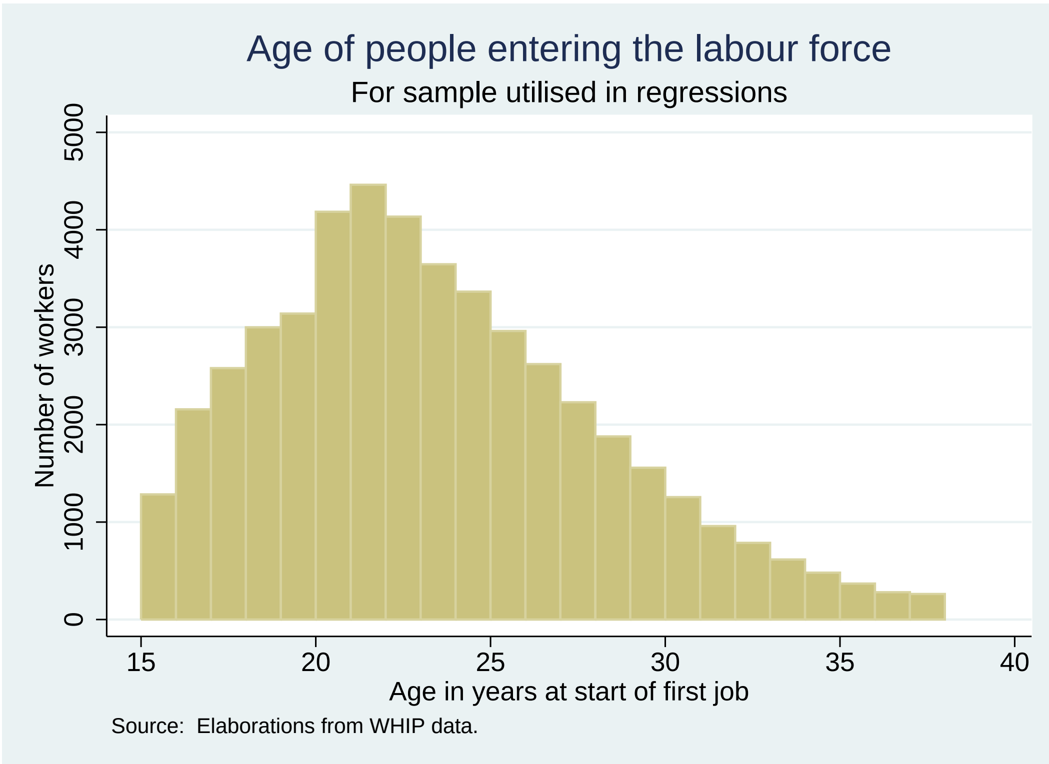

Figure 1: Distribution of workers by age at which they start their first job

Table 1: **Region of birth of workers**

Item	Frequencies	Percentages
North-West	12,033	24.9
North-East	8,872	18.3
Centre	7,775	16.1
South	10,222	21.1
Islands	4,924	10.2
Abroad	4,531	9.4
Total	48,357	100.0

Source: Elaborations from Whip Data

Table 2: **Region of Italy where jobs are located**

Item	Number	Per cent
North-West	14,727	30.4
North-East	11,672	24.1
Centre	9,425	19.4
South	8,355	17.2
Islands	4,286	8.8
Abroad	5	0.0
Total	48,470	100.0

Source: Elaborations from WHIP Data

Table 3: **Distribution of occupations for each workers' first job**

Item	Number	Per cent
Apprentice	10,152	20.9
Worker	25,599	52.8
Clerk	12,634	26.1
Manager	52	0.1
Director	30	0.1
Total	48,467	100.0

Source: Elaborations from Whip Data

Table 4: **Economic sector of the firm**

Item	Number	Per cent
Natural resource extraction	97	0.20
Manufacturing	18,714	38.73
Electricity and natural gas	130	0.27
Construction	5,722	11.84
Wholesale and retail trade, auto reparations	8,454	17.50
Hotels and restaurants	4,948	10.24
Transport, warehouses and communications	1,960	4.06
Banking and financial intermediaries	990	2.05
Computing, rental, research and other business sectors	5,476	11.33
Other Social and personal services	1,825	3.78
Total	48,316	100.00

Source: Elaborations from WHIP Data

Table 5: **Economic sector of the firm by Sex**

	Sex		
Economic sector of the firm - Ateco91	Female	Male	Total
Natural resource extraction	0.04	0.31	0.20
Manufacturing	36.08	40.50	38.73
Electricity and natural gas	0.18	0.33	0.27
Construction	1.76	18.57	11.84
Wholesale and retail trade, auto reparations	21.34	14.93	17.50
Hotels and restaurants	12.95	8.44	10.24
Transport, warehouses and communications	2.54	5.07	4.06
Banking and financial intermediaries	2.55	1.71	2.05
Computing, rental, research and other business sectors	16.31	8.01	11.33
Other Social and personal services	6.25	2.13	3.78
Total	100.00	100.00	100.00

Source: Elaborations from WHIP Dataset

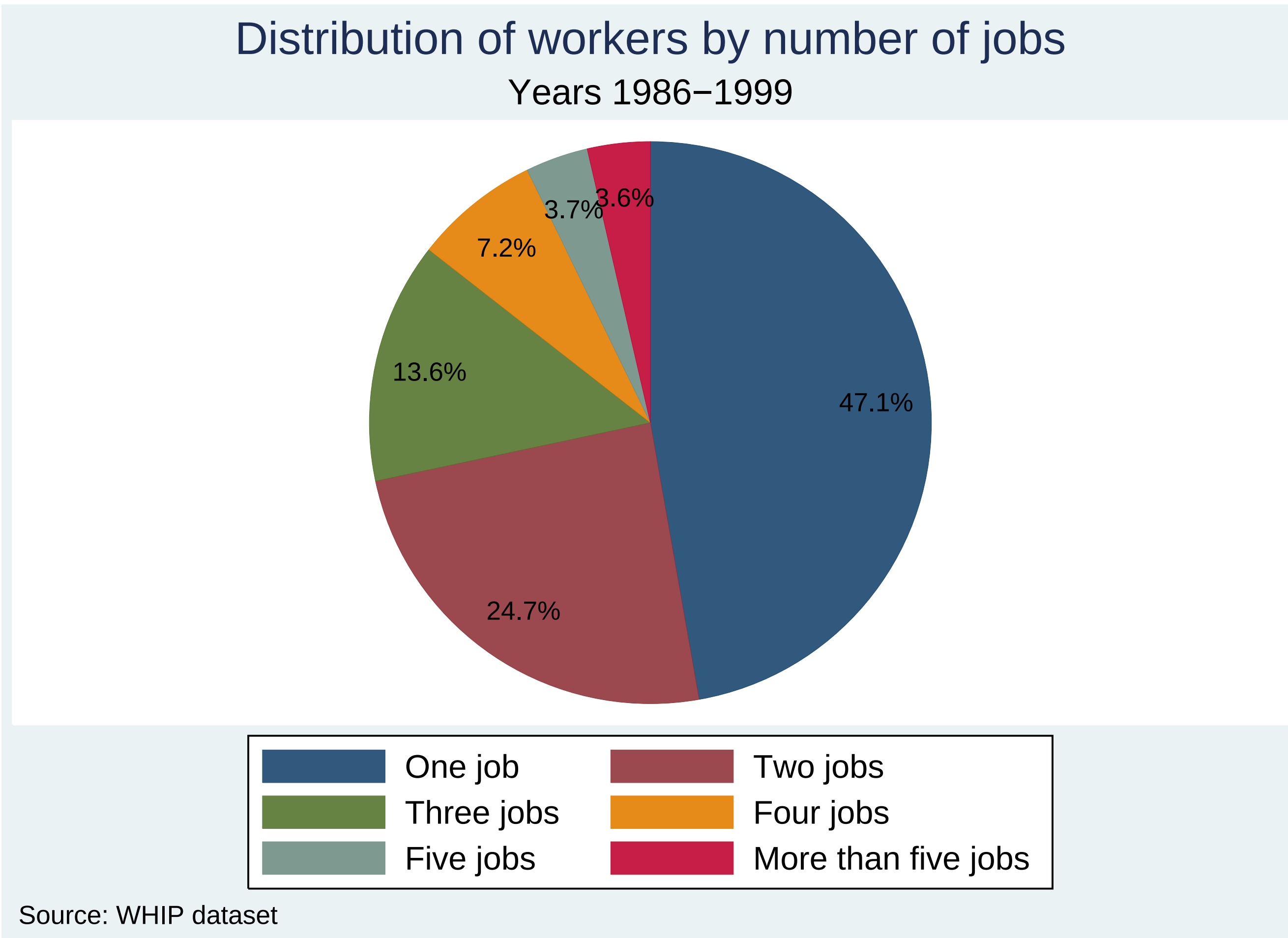

Figure 2: Distribution of workers by number of jobs

A.2 Jobs

Table 6: Summary statistics for Job Covariates

Variable	Mean	(Std. Dev.)	Min.	Max.	N
Dummy for Males	0.63	(0.48)	0	1	101921
Spell duration	1.87	(2.47)	0.08	13.91	101921
Dummy for censored spells	0.23	(0.42)	0	1	101921
Experience at beginning of spell (years)	1.11	(1.94)	0	12.93	101921

Table 7: Summary statistics for Job Covariates - Females only

Variable	Mean	(Std. Dev.)	Min.	Max.	N
Spell duration	1.98	(2.56)	0.08	13.91	37881
Dummy for censored spells	0.24	(0.43)	0	1	37881
Experience at beginning of spell (years)	1.06	(1.95)	0	12.93	37881

Table 8: Summary statistics for Job Covariates - Males only

Variable	Mean	(Std. Dev.)	Min.	Max.	N
Spell duration	1.81	(2.42)	0.08	13.91	64040
Dummy for censored spells	0.23	(0.42)	0	1	64040
Experience at beginning of spell (years)	1.14	(1.93)	0	12.88	64040

A.3 Annual

Table 9: Summary statistics for Year-level Covariates

Variable	Mean	(Std. Dev.)	Min.	Max.	N
Real yearly wages (FTE, in 1999 Euros)	16578.48	(7675.47)	0	99921.64	252594
Experience starting the year (in years)	2.68	(2.86)	0	12.93	252594
Tenure starting the year (in years)	1.64	(2.37)	0	12.93	252594

A.4 Earning Profiles

Figure 3: Experience Profile based on annual data

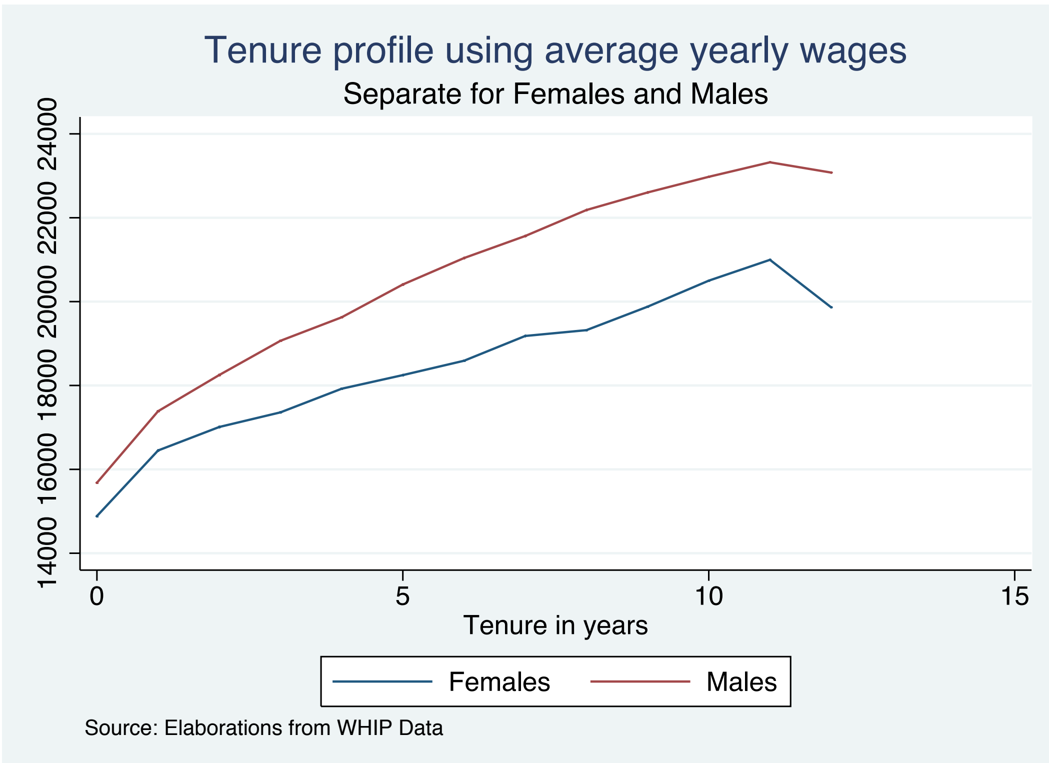

Figure 4: Tenure Profile based on annual data

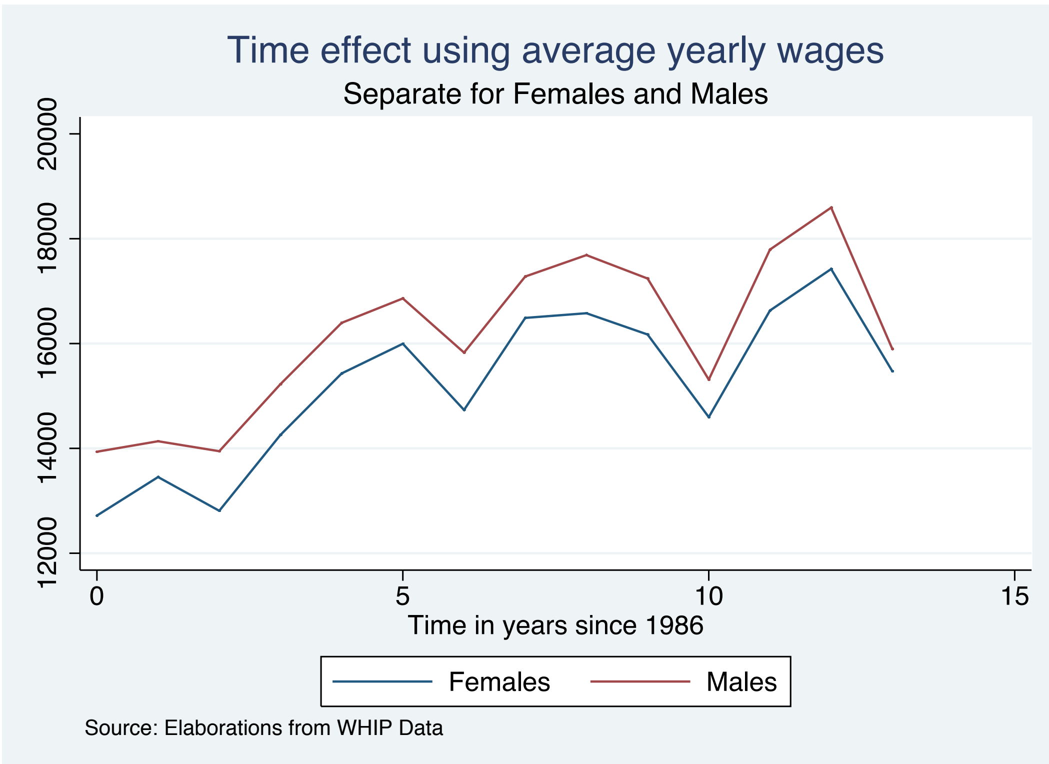

Figure 5: Time effect based on annual data

B Regression Tables

Table 10: Estimation results - Hazard Equation

Dependent variable: $ln(h_{iJ(i,t)t}(\tau))$			
	Models		
Variables	H1	H2	SM
Constant	-0.1831***	-0.4155***	-0.4257***
	(0.0100)	(0.0135)	(0.0137)
Male	0.0932***	0.1063***	0.1078***
	(0.0064)	(0.0103)	(0.0103)
Time trend	0.0227***	0.0182***	0.0187***
	(0.0010)	(0.0013)	(0.0013)
Seniority			
0-2nd year	-0.4243***	-0.1371***	-0.1291***
	(0.0071)	(0.0085)	(0.0086)
2nd-4th year	-0.0352***	0.0274***	0.0298***
	(0.0099)	(0.0104)	(0.0104)
4th-8th year	-0.1010***	-0.0543***	-0.0480***
	(0.0097)	(0.0103)	(0.0103)
8th year +	0.0091	0.0413*	0.0414*
	(0.0222)	(0.0230)	(0.230)
Experience			
0-4th year	-0.1834***	-0.2435***	-0.2426***
	(0.0039)	(0.0046)	(0.0046)
4th-8th year	-0.0555***	-0.0321***	-0.0362***
	(0.0064)	(0.0065)	(0.0066)
8th year +	-0.0718***	-0.0707***	-0.0705***
	(0.0145)	(0.0145)	(0.0145)
log L	-328098.81	-325956.61	-392220.75
$N = 101,917$			

H1: Hazard model without Individual Unobserved Heterogeneity

H2: Hazard model with Individual Unobserved Heterogeneity

SM: Simultaneous model using H2 specification for the job duration equation

Asymptotic Standard Errors in Parenthesis

Significance: *=10%; **=5%; ***=1%

Table 11: Estimation results - Wage Equation

Dependent variable: $ln(w_{iJ(i,t)t})$

	Models			
Variables	W1	W2	W3	SM
Constant	9.4376***	9.4233***	9.4265***	9.4266***
	(0.0028)	(0.0034)	(0.0035)	(0.0035)
Male	0.0428***	0.0468***	0.0469***	0.0469***
	(0.0025)	(0.0034)	(0.0033)	(0.0033)
Male × Exp.	0.3182***	0.2390***	0.2208***	0.2124***
	(0.0327)	(0.0281)	(0.0273)	(0.0271)
Time trend	0.0048***	0.0059***	0.0057***	0.0053***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Seniority				
0-2nd year	0.0288***	0.0262***	0.0263***	0.0213***
	(0.0015)	(0.0013)	(0.0012)	(0.0012)
2nd-4th year	-0.0020	-0.0123***	-0.0159***	-0.0167***
	(0.0021)	(0.0020)	(0.0016)	(0.0016)
4th-8th year	0.0165***	0.0079***	0.0043***	0.0036***
	(0.0022)	(0.0016)	(0.0011)	(0.0011)
8th year +	0.0043	0.0008	-0.0002	-0.0009
	(0.0047)	(0.0032)	(0.0021)	(0.0021)
Experience				
0-4th year	0.0399***	0.0467***	0.0451***	0.0450***
	(0.0011)	(0.0012)	(0.0012)	(0.0012)
4th-8th year	0.0072***	0.0071***	0.0083***	0.0088***
	(0.0013)	(0.0005)	(0.0004)	(0.0004)
8th year +	0.0146***	0.0124***	0.0133***	0.0137***
	(0.0027)	(0.0011)	(0.0008)	(0.0008)
log L	-77705.74	-68860.01	-66459.90	-392220.75

$N = 252,586$

W1: Wage model without Individual Unobserved Heterogeneity

W2: Wage model with Individual Unobserved Heterogeneity

W3: Wage model with Individual and Match Unobserved Heterogeneity

SM: Simultaneous model using W3 specification for the job duration equation

Asymptotic Standard Errors in Parenthesis

Significance: *=10%; **=5%; ***=1%

Table 12: Estimation results - Estimates for Variance components and Parameters

	Models			
	H1+W1	H1+W2	H2+W3	SM
Individual Heterogeneity variance components				
σ_{θ_1}			0.6713***	0.6680***
			(0.0072)	(0.0074)
σ_{θ_2}		0.2819***	0.2626***	0.2621***
		(0.0006)	(0.0007)	(0.0007)
σ_{θ_3}		1.0938***	1.1696***	1.1711***
		(0.0289)	(0.0306)	(0.0308)
$\rho_{\theta_2\theta_3}$		-0.4834***	-0.5513***	-0.5538***
		(0.0057)	(0.0058)	(0.0058)
$\rho_{\theta_1\theta_2}$				0.0126
				(0.0113)
$\rho_{\theta_1\theta_3}$				-0.0115
				(0.0203)
Match Heterogeneity variance components				
σ_δ			0.1946***	0.1954***
			(0.0004)	(0.0005)
ϕ				-0.7339***
				(0.0378)
Error structure				
ω	0.6654***	0.3720***	0.1075***	0.1057***
	(0.0004)	(0.0008)	(0.0015)	(0.0015)
σ_ν	0.2925***	0.2691***	0.2399***	0.2397***
	(0.0001)	(0.0001)	(0.0002)	(0.0002)

H1+W1: no unobserved heterogeneity components
H1+W2: Individual Unobserved Heterogeneity in the Wage equation only
H2+W3: Hazard model with Individual Unobserved Heterogeneity, Wage equation with Individual and Match Unobserved Heterogeneity Components
SM: Simultaneous model

Asymptotic Standard Errors in Parenthesis
Significance: *=10%; **=5%; ***=1%

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), p251 – 333.

Altonji, J. G. and R. A. Shakotko (1987). Do wages rise with job seniority? *Review of Economic Studies 54*(3), p437 – 459.

Becker, G. (1964). *Human Capital.* The University of Chicago Press.

Brown, J. N. (1989). Why do wages increase with tenure? on-the-job training and life-cycle wage growth observed within firms. *American Economic Review 79*(5), p971 – 991.

Contini, B. (2002). *Osservatorio sulla Mobilità del Lavoro in Italia.* Il Mulino.

Contini, B., R. Leombruni, L. Pacelli, and C. Villosio (2007). Wage mobility and dynamics in italy in the 90's. *NBER Working Paper No. 13029*.

Contini, B. and U. Trivellato (2005). *Eppur si muove. Dinamiche e Persistenze nelo Mercato del Lavoro Italiano*. Bologna: Il Mulino.

Dohmen, T. J. (2004). Performance, seniority, and wages: Formal salary systems and individual earnings profiles. *Labour Economics 11*(6), p741 – 763.

Dostie, B. (2005). Job turnover and the returns to seniority. *Journal of Business and Economic Statistics 23*(2), p192 – 199.

Dustmann, C. and C. Meghir (2005). Wages, experience and seniority. *Review of Economic Studies 72*(1), p77 – 108.

Gibbons, R., L. F. Katz, T. Lemieux, and D. Parent (2005). Comparative advantage, learning, and sectoral wage determination. *Journal of Labor Economics 23*(4), p681 – 723.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy 87*(5), p972 – 990.

Kiefer, N. M. (1988). Economic duration data and hazard functions. *Journal of Economic Literature 26*(2), p646 – 679.

Lillard, L. A. (1999). Job turnover heterogeneity and person-job-specific time-series wages. *Annales d'Economie et de Statistique* (55-56), p183 – 210.

Lillard, L. A. and C. Panis (2003a). aml multilevel multiprocess statistical software, release 2.0.

Lillard, L. A. and C. Panis (2003b). *aML User Guide and Reference Manual*. EconWare, Los Angeles, California.

Nagypal, E. (2007). Learning by doing vs. learning about match quality: Can we tell them apart? *Review of Economic Studies 74*(2), p537 – 566.

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of Labor Economics 13*(4), p653 – 677.

Parent, D. (2000). Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics. *Journal of Labor Economics 18*(2), p306 – 323.

Pissarides, C. A. (1994). Search unemployment with on-the-job search. *Review of Economic Studies 61*(3), p457 – 475.

Poletaev, M. and C. Robinson (2008). Human capital specificity: Evidence from the dictionary of occupational titles and displaced workers surveys, 1984-2000. *Journal of Labor Economics 26*(3), 387–420.

Pollard, J. and E. Valkovics (1992). The gompertz distribution and its applications. *Genus*.

Topel, R. H. (1991). Specific capital, mobility, and wages: Wages rise with job seniority. *Journal of Political Economy 99*(1), p145 – 176.