WARWICK
THE UNIVERSITY OF WARWICK

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:
http://wrap.warwick.ac.uk/109758

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

THE BRITISH LIBRARY
BRITISH THESIS SERVICE

COPYRIGHT

Reproduction of this thesis, other than as permitted under the United Kingdom Copyright Designs and Patents Act 1988, or under specific agreement with the copyright holder, is prohibited.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

REPRODUCTION QUALITY NOTICE

The quality of this reproduction is dependent upon the quality of the original thesis. Whilst every effort has been made to ensure the highest quality of reproduction, some pages which contain small or poor printing may not reproduce well.

Previously copyrighted material (journal articles, published texts etc.) is not reproduced.

THIS THESIS HAS BEEN REPRODUCED EXACTLY AS RECEIVED

On the Covariance Structure and Mobility of Italian Wages

by

Lorenzo Cappellari

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Economics

University of Warwick, Department of Economics

September 1999

to Anni

Table of contents

List of tables and illustrations

Acknowledgements

I am deeply grateful to my supervisor Mark Stewart, who guided me throughout these three years of research work, teaching me a lot and making econometrics look easier. Working with him has been an invaluable experience.

I owe a debt of gratitude to Carlo Dell'Aringa e Claudio Lucifora, who, after introducing me to this research area, advised and encouraged me over the years, representing an essential point of reference for my work.

I would like to express appreciation to my Ph.D examiners Wiji Arulampalam and Stephen Jenkins for their enlightening remarks on the contents of this Thesis.

During these years I have benefited of comments and conversations with participants at seminars and conferences; among these, I especially thank Paul Bingley, Andrea Ichino, Markus Jäntti, Francis Kramarz and Andrew Oswald. I also thank an anonymous referee for remarks on Chapter 2.

Financial assistance from the Università di Pavia, the Università Cattolica di Milano and the ESRC (award no. R00429724408) and data availability from the Italian Social Security Institute and the Bank of Italy, the latter through IGIER, are gratefully acknowledged.

Of course, ideas and errors remain my sole responsibility.

The support of Annalia, Maria Rosa, Pietro and Tiziano made everything possible.

Declaration

This Thesis is my own work. A paper based on Chapter 2, which has originally been written for this Thesis, has been accepted for publication in *The Manchester School.* The material of this Thesis has not been submitted for a degree elsewhere.

Abstract

This Thesis uses Italian panel micro-data to investigate the intertemporal wage covariance structure and the extent of transition probabilities at the bottom of the wage distribution, producing new and original evidence on the degree of persistence of cross-sectional wage differentials over individual life-cycles and on the features of the process governing wage mobility across low-pay thresholds. Chapter 1 presents a survey of the debate generated by the rise of wage inequality observed in many industrialised economies and stresses how longitudinal analyses of wage persistence and mobility shed light on the long term impact of rising cross-sectional dispersion; a survey of the two research areas to which this Thesis contributes, i.e. variance components models of the wage covariance structure and econometric modelling of transition probabilities, is also presented.

Variance components models of the wage covariance structure are estimated in Chapters 2 and 3, where two unbalanced panels drawn from the Social Security archive on the 1974-88 and 1979-95 intervals (respectively) are analysed by applying the minimum distance technique. Chapter 2 shows that while permanent wage profiles converged within the overall wage distribution, divergence can be detected for white collar workers, suggesting that the former could have been imparted by the egalitarian wage policies of the late 1970s. Results from Chapter 3 indicate that the rising wage inequality observed in Italy over the 1980s and the early 1990s permanently affected the evolution of wage profiles especially during the second half of the 1980s; on the other hand, increases in the relative importance of wage volatility are shown to characterise the first half of the 1990s, thus mirroring the higher labour market "flexibility" of recent years. The Chapter also takes into account the relationship between covariance structure components and observable workers characteristics; in particular, a model which shifts the parameters of interest with respect to workers' occupations is developed, finding that permanent differentials arise from the wage distribution of white collar workers.

A bivariate probit model with endogenous switching is developed in Chapter 4 to analyse low-wage mobility taking the endogeneity of starting wage states into account, using survey data from the Bank of Italy. Results indicate the appropriateness of such a framework, the correlation between state and transition probabilities being statistically significant. While workers' attributes are found to have a limited impact on the probability of leaving low-pay, a considerable share of aggregate low-pay persistence appears to be the consequence of *true state dependence*, i.e. the experience of low-pay raises, *per se*, the likelihood that the phenomenon occur in the future. Chapter 5 checks the robustness of these conclusions to the presence of endogenous attrition from the wage distribution over time by augmenting the model with a third equation for the probability of belonging to the balanced sample. The computational difficulties posed by the required evaluation of trivariate normal integrals are overcome by implementing simulation estimation techniques. Results indicate that exits from the wage distribution over time are an ignorable source of sample selection for the estimation of low-pay transition probabilities on these data, thus pointing towards the robustness of the findings of Chapter 4 to this generalization of the model.

Introduction

Increasing levels of wage inequality have characterised the labour market in many industrialised economies over the recent past, a stylised fact which has become of central relevance for the debate in labour economics. A large body of literature[1] documents how cross-sectional wage differentials have been growing in several directions and, as pointed out in these studies, positive and normative considerations related to this phenomenon are linked to each other. Under the first viewpoint, understanding which forces are driving the rise in inequality can shed light on labour market functioning and the way wage determination mechanisms react to structural changes in the labour market. On the other hand, disentangling the relationship between wage inequality and workers' characteristics has relevant policy implications, since it can indicate the source of rising inequality and what kind of intervention is needed to prevent it from growing further.

Despite being traditionally known as "rigid", the Italian labour market shares this tendency of widening gaps between the top and the bottom of the wage distribution. Several studies[2] have shown how, after dropping considerably over the late 1970s and early 1980s, measures of inequality tend to rise sharply during the second half of the 1980s and the early 1990s. These dynamics paralleled the remarkable changes in the wage setting framework taking place during the 1980s. As an example, in 1981 the system of wage indexation to the cost of living foresaw equal monetary increases all over the wage distribution, while ten years later any form of automatic wage protection against inflation had been abolished.

To observe that wage differentials measured on cross-sectional data grow over time is only partially informative about the degree of inequality generated by the labour market. Is the increasing dispersion persistently affecting workers' careers or

[1] These studies are surveyed in Section 1.2.
[2] See Section 1.3 for a review of this literature.

is it just the outcome of wage volatility with transient effects? Is the wage rank of an individual in a given period a good predictor of her future ranks or is it likely to change? On the one hand, individuals may move within the classes of the wage distribution through time, so that those who occupy the lowest quantiles in a given year will not necessarily be observed in the same hierarchical position after a few periods. In such a case, growing cross-sectional dispersion doesn't entirely translate into higher inequality in the medium or long run at the individual level: the labour market provides a dynamic redistribution of labour incomes and low-wage experiences are shared among individuals through their careers. Conversely, even if cross-sectional inequality is constant, the distribution could be characterised by a high degree of hierarchical persistence, so that individuals observed at the bottom end at one point in time have a high probability of also experiencing low-wages in the future. In this event, the labour market would produce inequality in a dynamic sense and cross-sectional measures of wage dispersion would not pick-up such a phenomenon. To answer the questions above and gain a complete picture of the wage distribution's dynamics it is necessary to follow individual wage profiles over time, assessing the persistence of cross-sectional differentials and the probability of transitions within the classes of the distribution through time.

The multi-period perspective offered by the analysis of wage persistence and mobility can improve our understanding of the wage inequality issue from both the positive and normative viewpoint. In the first case, the distinction between permanent and transitory components of wage inequality can indicate which is the nature of the structural labour market changes causing the dynamics of aggregate differentials, as long as different kinds of shocks impact on each component with different intensities. For what concerns the policy implications, while assessing the degree of hierarchical persistence yields indications about the need of interventions to protect the low-paid, disentangling the workers' attributes generating transitions out of low-pay is informative about the kind of policies which could help in fighting long-term labour market poverty.

Studies of persistence and mobility require the availability of panel data on individual wages which make it possible to move from the analysis of the moments of the marginal wage distribution (typical of the wage inequality literature) to the assessment of those of the joint distribution over different points in time. The availability of multiple wage observations for the same individual over time allows identification of the long run wage component, so that a distinction can be made between permanent and transitory components of aggregate cross-sectional differentials. Also, panel data can be used to track individual movements within the classes of the wage distribution through time, so that the probability of changes in wage ranks can be estimated.

This Thesis uses Italian panel micro-data to investigate the degree of persistence and mobility of the wage distribution. While a growing body of literature analyses cross-sectional wage differentials in Italy, longitudinal studies are still rare and both the use of panel data and the methodological approaches followed in this study will provide a substantial improvement of our knowledge of the interplay between individual wage dynamics and wage inequality. Chapter 1 introduces the theme of wage persistence and mobility, stressing the relevance of this kind of analyses for the wage inequality debate. After reviewing the main points that have emerged from the wide debate on wage inequality and the relevant facts on the Italian wage distribution, the Chapter shows how cross-sectional inequality and intertemporal persistence (immobility) are the two components of long term inequality, and how analyses of mobility are relevant in various circumstances usually deemed to depend upon cross-sectional wage differentials, such as human capital investments or the design of low-wage protection policies. The literature on wage persistence and mobility is then reviewed, distinguishing between the two research areas to which this Thesis will contribute: variance components models of the intertemporal wage covariance structure and the econometric modelling of wage transition probabilities.

Chapter 2 analyses the intertemporal male wage covariance structure using administrative unbalanced panel data for the 1974-1988 interval, the data set being the only one available from this source at the time this research began. Variance components models which allow an assessment of the degree of persistence and volatility within the second moments of the joint intertemporal distribution of wages are estimated by applying the minimum distance technique, i.e. the restrictions on the theoretical covariance structure implied by the model are imposed directly on empirical second moments. A central issue addressed in this Chapter is the extent of convergence in permanent wages, which is assessed by specifying them as individual linear profiles (so called random growth models) whose intercepts and slopes are allowed to covary over individuals. Such a specification has been proposed to test human capital theories of wage dynamics, a negative covariance between individual intercepts and slopes being in line with the catching up of wage profiles predicted by the Mincerian on-the-job training version of the theory. Moreover, cross-overs of permanent wage profiles imply that wage mobility is generated by permanent workers' characteristics rather than by volatility, thus preventing welfare worsening effects induced by wage uncertainty. Transitory wages are specified according to low-order ARMA processes which allow for the fact that, contrary to time series analysis, in a panel micro-data context initial conditions of the process cannot be placed in the infinite past and need to be explicitly modelled.

Estimated models are also characterised by polynomial loading factors on each wage components, aimed at controlling for time-wise heterogeneity in the covariance structure. Finally, additional features of individual wage dynamics are addressed by performing the analysis within samples defined by workers' occupation.

The analysis of male wage persistence and volatility is extended along several directions in Chapter 3. First of all, the extent of time-wise heterogeneity in the wage covariance structure is assessed by means of period-specific shifters which do not impose any functional form, so that the relative importance of changes in permanent wage differentials and wage volatility in determining aggregate inequality dynamics can be flexibly estimated. A second relevant feature of the analysis in this Chapter is the comparison between alternative dynamic specifications of the permanent wage, namely random growth versus random walk models, the latter being aimed at capturing wage persistence through the unit root hypothesis. Thirdly, the relationship between variance components and workers attributes is analysed in detail and a model which shifts covariance structure parameters with respect to individuals' characteristics proposed. Finally, a larger and more recent administrative panel is analysed, thus providing evidence also for the years in which automatic wage indexation has been completely abolished.

The issue of hierarchical mobility at the bottom of the wage distribution is addressed in Chapter 4, where an econometric model of low-wage transition probabilities is estimated. The perspective proposed in this Chapter is alternative and complementary to the one of variance components models: while in that case the whole distribution is subsumed under a limited number of parameters and the effect of both observable and unobservable workers' characteristics on persistence is taken into account within the permanent wage component, the low-wage mobility model concentrates on wage dynamics at the bottom of the distribution and estimates the effect of each observable attribute on the probability of crossing pre-defined low-wage thresholds. The modelling strategy follows and extends the framework proposed by Stewart and Swaffield [1999] who use bivariate probits with endogenous switching to analyse low-pay transitions controlling for the endogeneity of the conditioning starting state, the so called initial conditions problem. Implementation of this model is demanding in terms of data, since the availability of instruments to identify the initial conditions of the transition is required. Household survey data containing information preceding the labour market entry are used. However, such information is available only for recent waves of the survey, so that the analysis focuses on the 1993-95 transition; moreover, due to the small sample size, female and male data are analysed jointly. After estimating the effect of observable characteristics on low-wage transitions, the model's predictions are used to disentangle the extent of pure state dependence on aggregate persistence, thus shedding light on the role played by the experience of low-pay itself in determining the occurrence of low-pay episodes in the future.

Further insights on low-pay transitions are provided in Chapter 5, where the robustness of the results previously obtained to the presence of endogenous sample selection is assessed. Estimation of the model in Chapter 4 is based on a balanced sample, i.e. only observations with a valid wage at both ends of the transitions are used, while those exiting the group of wage earners are discarded from the analysis. As long as the process governing attrition from the wage distribution is correlated with low-pay transitions, such a selection rule may lead to biased estimates. To cope with this issue, an equation for the probability of staying in the sample and observations on those leaving the group of wage earners during the transition are added to the model, yielding a trivariate probit. However, estimation of the model poses a computational hurdle, since trivariate normal integrals are not packaged within statistical software. To overcome the problem, the method of simulated maximum likelihood has been implemented, in which the intractable bit of the objective function is replaced by its simulated counterpart; details on the implementation of the simulated maximum likelihood estimator are given in the Appendix to the Chapter. Simulated estimation of this model will thus enable us to set-up a more general framework for the analysis of low-pay transition probabilities, within which the potential endogeneity of panel attrition, and hence the robustness of results provided in Chapter 4, can be evaluated.

Chapter 1

Individual wage dynamics, persistent inequality and mobility

1.1 Introduction

This Chapter presents a survey of the literature on wage persistence and mobility and stresses the relevance of this kind of analysis for the more general debate on wage differentials and wage inequality. The study of wage inequality has become increasingly popular in recent years as an attempt to provide explanations for the considerable changes in the wage distribution observed in many industrialised economies. This literature typically utilises cross-sectional data and studies variations in the marginal wage distribution, i.e. the distribution of wages at a given point in time; for this reason it misses important dynamic aspects of the changes in wage inequality which can be assessed by utilising panel data to analyse wage persistence and mobility. The use of panel data in which the same worker is observed over several years allows the adoption of a different perspective: the joint wage distribution over various points in time can be studied, so that not only the extent, but also the persistence of wage differentials through time can be analysed. Such a different perspective generates an informative gain both for our understanding of labour market functioning and for the design of policies aimed at coping with increasing wage inequality.

The Chapter is organised as follows. Section 2 offers a survey of the debate raised by the changes in the wage distribution characterising several countries over the last two decades. The survey is not meant to be exhaustive, but is aimed at highlighting the main points emerged from such a wide debate. Relevance is given to the stylised facts shaping the debate, with particular reference to the United States, for which distributional dynamics have been extensively documented, and with some element of the European experience. The Section then develops by illustrating the main theoretical explanations put forward to account for the empirical evidence. Section 3 illustrates the stylised facts on the Italian wage distribution, stressing how institutional developments in the labour market have been relevant in determining the evolution of wage differentials. The link between cross-sectional wage inequality, persistence and mobility is introduced in the fourth Section, where it is shown how longitudinal data allow the adoption of a multi-period approach to inequality, which is relevant under both the positive and normative viewpoint. The last two Sections give an illustration of the empirical literature on wage persistence and mobility, making a distinction between the two main bodies of research to which this Thesis will contribute: the GMM analysis of the wage intertemporal covariance structure and the estimation of models for the probabilities of transition within the classes of the wage distribution through time. The line of argument is, in these two Sections, descriptive; the formal discussion will be developed in the next Chapters.

1.2 Wage inequality: stylised facts and proposed explanations

The marked growth of wage differentials undoubtedly represents a central phenomenon characterising the labour market of many industrialised countries over the last two decades. Its importance is mirrored by the large body of, mainly empirical, research dealing with the dynamics of the wage distribution in recent years; quoting the classical survey paper of Levy and Murnane [1992]:

"...*within a decade, earnings inequality grew from a lightly studied branch of labor economics to a major research area...(pag. 1334)*".

The analysis of the features characterising this process of growing differentials and of the causes behind it is relevant both for the understanding of the mechanisms of wage determination and for the study of policy measures aimed at coping with the welfare consequences of growing inequality. Under the first viewpoint, researchers have focused both on variations in the relative supply and demand for skilled labour and on the impact of developments in labour market institutions, and the interaction between such explanations has generated a paradigm within which an organic view of structural labour market changes can be gained. On the other hand, the welfare implications of growing inequality arise from the fact that it typically implies a loss in the relative purchasing power of workers at the bottom of the wage distribution (OECD [1996]), and that a growing number of individuals earns wages below fixed "decency thresholds". In order to properly design policy interventions to fight labour market poverty, it is then important to understand the causes generating such losses and to identify the characteristics of the individuals who find themselves worse-off.

1.2.1 Stylised facts

A considerable share of the empirical evidence on growing wage inequality refers to the US labour market, a fact which probably reflects both the earlier emergence of the phenomenon and its higher intensity compared to Europe. The features of the US rise in wage inequality are thus well documented and, in particular, the following stylised facts have shaped the debate:

- a growth in the dispersion of raw labour incomes can be observed since the first half of the 1970s, with an acceleration during the 1980s decade (Levy and Murnane [1992]);
- wage differentials by education follow a sort of sine wave path, with a peak located around 1970, followed by a drop over the whole 1970s decade and a drastic re-opening during the 1980s (Murphy and Welch [1992]);
- wage differentials by years of labour market experience grow for each educational level during the 1970s, while during the 1980s they kept on

growing only for the less educated segment of the labour force (Katz and Murphy [1992]);

- since the early 1970s wage dispersion grew steadily also within the cells defined by educational levels and labour market experience (Juhn et al. [1993]);
- the only differential which dropped consistently between the 1970s and the 1990s was the gender wage gap (Gottschalk [1997]).

Juhn et al. [1993] analyse the effects of changes in labour force composition, changes in "prices" for such characteristics (i.e. estimated coefficients in wage regressions) and changes in residual (within-groups) variability on wage dispersion at various quantiles of the distribution. Using CPS data, these authors show that while the dynamics of personal characteristics had a limited impact on overall inequality, the evolution of their prices had a strong effect in the upper part of the distribution, especially during the 1980s and, on the other hand, within-groups variation accounted for relative wage losses in the bottom quantiles. Gottschalk and Moffitt [1994] use PSID panel data to compute medium term wage measures and the yearly deviations from them; they show that a relevant share (approximately one third) of growing cross-sectional dispersion can be ascribed to increasing wage instability.[3]

The dynamics of the wage distribution has typically generated increases in wage dispersion also in European labour markets. As documented in several issues of the OECD Employment Outlook [1993, 1996], the intensity of such changes is rather heterogeneous. Great Britain is the European country experiencing by far the

[3] Being based on panel data this study deals with individual wage dynamics and the issue of changes in medium term inequality. However, it is reviewed here, rather than in the Sections centred on models of long term inequality and mobility, given its relevant impact on the debate on inequality.

strongest growth in wage inequality, especially over the 1980s. Gosling et al. [1998] show how, during the 1980s and the early 1990s wage growth was rather limited for workers at the bottom of the distribution, while at the median and, especially, for the highest quantiles, it was much faster, generating, consequently, a drastic reopening of differentials. Stewart and Swaffield [1997] report an increased incidence of low-wage jobs since the mid-1980s.

Taking other European countries into account, it has been shown how the growth in measures of wage inequality has been less intense and/or has affected only certain dimensions of wage differentials (OECD [1993]).[4] Typically, the existing literature suggests that, among those countries for which some evidence has been produced, the only one which didn't experience any re-opening of differentials is the former West Germany (Gottschalk [1997]; Gottschalk and Smeeding [1997]). Abraham and Houseman [1993] report steady levels of wage dispersion over the 1980s and suggest that the combination of the efficient German system of education, capable of satisfying the demand for various types of skilled labour, and of the highly unionised system of wage bargaining have been the main factors behind the distribution's stability.

1.2.2 Proposed explanations

As anticipated above, the search for explanations capable of accounting for the empirical evidence has been mainly based on schemes of labour supply and demand. In particular, given that inequality grew along both observable (education and experience differentials) and unobservable (within-groups variance) dimensions of skills, rising differentials have been in many circumstances associated with a rise in the relative "price" for skilled labour and researchers' attention has been focused

[4] The discussion on the dynamics of the Italian wage distribution is postponed to the next Section.

on the identification of structural changes in the labour market, especially in the US one, likely to alter the balance between demand and supply of skills. In parallel, alternative explanations emphasising the decline in labour market institutions have been put forward.

The evidence of wage differentials rising by degree of education or labour market experience requires the presence of factors capable of raising the relative demand for such workers characteristics. An explanation frequently advanced by the literature identifies such a factor in the non-neutrality of technological change with respect to different types of labour, in particular skilled versus unskilled, an hypothesis which is widely known as *skilled biased technical change*. According to this explanation, the introduction of production techniques with high technological content (such as the diffusion of personal computers in the workplace) requires an increase in the skill endowment of the labour force and, holding the supply of such skills fixed, alters the relative wage in favour of more skilled workers (Bound and Johnson [1992]; Johnson [1997]). In practice, this hypothesis assumes that the growth in the demand for skills has been faster than that in supply; evidence supporting this view is reported in Berman et al. [1997]. It been observed how, similarly to Solowian theories of economic growth, the hypothesis treats technical progress as a residual: given that technical change its hardly observable, its relevance would result from the inability to find alternative factors capable of explaining the empirical evidence. However, some attempts have been made to directly measure the impact of technical change on wages. In particular, evidence in favour of the existence of a wage premia for the use of personal computers has been produced for the US and the UK (see Krueger [1993] and Bell [1996], respectively), while opposite results have been provided for France (Entorf and Kramarz [1997]).

The main alternative market-based explanation for the rise in wage inequality is centred on the impact of international trade with less developed countries. According to this line of argument, the increasing degree of openness towards countries producing goods with a low skill content has put less skilled workers of industrialised economies in competition with similar workers in developing countries, thus depressing their wages (Wood [1995]). Evidence from macro-data supporting this hypothesis has been provided by Borjas and Ramey [1994]. The main critique to this hypothesis stresses how wage dispersion has been growing also in sectors excluded from the globalisation process.

Some authors have stressed how also changes in labour supply could have played a role in generating increasing differentials (Topel [1997]). In this respect, the two major shocks on the supply side have been the entry of the baby-boom generation into the labour market between the 1960s and the 1970s and the increasing female labour force participation. The first phenomenon can account for the development of differentials for higher education and labour market experience during the 1970s (Katz and Murphy [1992]); the baby-boom wave has altered the composition of the labour force, which has become younger and more educated, thus depressing wage rates in these directions. For what concerns female labour supply, the existence of gender discrimination in the labour market could imply that female workers entering the labour market will compete with less qualified male labour force (Topel [1994]): this would explain both the relative loss for males at the bottom of the distribution and the reduction in the *ceteris paribus* gender wage gap.

An explanation based on changes in the institutional environment arise from the observation that the rise in inequality has paralleled a generalised reduction in the degree of unionisation and, consequently, in trade unions' bargaining power. In general, the presence of unions can affect wage dispersion in two directions. On the one hand a positive differential arises for workers covered by contracts negotiated by unions (higher between groups inequality); such an effect is weaker the more these contracts are extended also to non unionised workers. On the other hand, the wage distribution for unionised workers is more compressed than it would be in the absence of unions (lower within groups inequality). The empirical evidence reported in Fortin and Lemieux [1997] for the US suggests a quantitative predominance of the second effect, so that de-unionisation has a net positive impact on wage dispersion. Evidence in favour of this hypothesis has been provided also by Gosling and Machin [1995] for the UK. More in general, the presence of stronger labour market institutions in continental Europe could account for the weaker growth of inequality compared to the UK and the US, holding fixed the intensity of technical change or international trade; as noticed by Blau and Kahn [1996], even if egalitarian wage policies, in particular aimed at protecting workers in the lowest quantiles of the wage distribution, have been in place also in the US, European unions have been more successful both in compressing wage differentials for unionised workers and in extending contract provisions to non-unionised workers.

1.3 Evidence on the dynamics of the Italian wage distribution

The relationship between changes in labour market institutions and the dynamics of wage differentials is particularly relevant in the Italian case. Wage egalitarianism has been one of the main goals on the unions' side during the 1970s and, as Accornero [1992] points out, it was also meant to maintain the high level of social support which unions enjoyed after the "hot Autumn" of 1969, a period of intense social conflict. The era of wage egalitarianism began in 1969 with the introduction of automatic fixed amount wage increments to compensate for inflation in the contract of metal manufacturing workers and reached its peak in 1977 when, following collective agreements between workers' and employers' confederations dating back to 1975, such fixed amount compensations (the so called *Scala Mobile a punto unico*)[5] were extended by law to all workers (Treu [1984]).[6] As we will see below, the egalitarian system of indexation imparted a drastic compression to wage differentials, which turned out to be hardly tolerable to white collar workers in the top part of the wage distribution. These workers progressively lost their attachment to unions and in 1980 a mass demonstration against egalitarianism (the *March of the 40,000*) was organised in Turin, marking a turning point in the history of industrial relations. The 1980s have witnessed a series of reforms of the wage indexation system towards proportionality (Erickson and Ichino [1995]), while in 1993 collective agreements were signed with which any form of automatic wage indexation was abolished. In parallel, other reforms, typically orientated towards a higher "flexibility" in industrial relations, have been introduced (Bertola and Ichino [1995]); the literature reviewed in this Section shows how, during the same period, the wage distribution was characterised by a reopening of differentials.

International comparisons of wage inequality usually rank Italy among those economies with a modest growth of differentials during the 1980s. The 1993 OECD Employment Outlook reports the measures of dispersion computed by Erickson and Ichino [1995] using survey data from the Bank of Italy and assigns a 0 (no variation) to the dynamics of the distribution over the 1980s (the period actually considered is 1978-1987). The 1996 edition of the same report adds three cross-sections from the

[5] The expression *Scala Mobile* (escalator) designates the system of wage indexation; the *punto unico* (single point) version is the one with monetary increases equal over the whole distribution.

[6] It has to be stressed how both unions (CGIL, CISL and UIL) and the confederation of employers (Confindustria) took advantage from the egalitarian reform of the escalator. If, in the first case, this constituted a "victory", in the second it amounted at introducing some additional rigidity in the process of wage determination which could help in containing strong calls for wage increases at the firm level originating from the high bargaining power of unions in the workplace (Accornero [1992]).

same survey to the data set (1989, '91 and '93), showing how a drastic increase of differentials characterises the beginning of the 1990s. In their survey paper of 1997, Gottschalk and Smeeding report the statistics of Erickson and Ichino, concluding that, together with the former West Germany, Italy

"...form(s) a small group that experienced no measurable increase in earnings inequality during the 1980s...(pag. 652)".

In his 1997 paper, Gottschalk reports evidence from the 1996 OECD Employment Outlook and ranks Italy among countries with moderate increases in wage dispersion, stressing how the bulk of growing differentials is concentrated at the beginning of the 1990s. In a recent international comparison Bardone et al. [1998] compare micro-data from the Bank of Italy to those from the National Social Security Institute[7] and show that, in the second case, the impression of growing inequality over the 1980s and early 1990s is clearer, both for intensity and timing; under this last viewpoint, while data from the Bank of Italy display growing differentials only from the beginning of the 1990s, Social security records show growing differentials since 1986, the first year of data available.

Apart from these international comparisons, a growing literature deals with the dynamics of the Italian wage distribution. The fundamental stylised fact reported is that the phase of distributional compression imparted by the egalitarian wage indexation system has been followed by a marked reopening of differentials.

Erickson and Ichino [1995] use micro-data from the Bank of Italy in conjunction, among others, with firm level data aggregated by occupation provided

[7] Both data sets will be used in this Thesis: their features are described in the next Chapters.

by local and sectoral confederations of employers (Assolombarda and Federmeccanica, respectively). The authors reach contrasting conclusions about the timing of the turning point in inequality dynamics depending upon the data source utilised. In particular, while cell data present a minimum in inequality series within the first half of the 1980s, individual data provide the picture of continuous drop over the 1980s already mentioned above. The information available in aggregated data enable them to decompose wage growth into the part bargained in national contracts (which includes Scala Mobile payments) and the one negotiated locally (the so called wage drift, determined at the firm level), showing how this last component has been used to neutralise the compression induced at the national level, with its effect prevailing since the mid-1980s. Moreover, they show how the tendency towards increasing educational wage premia has been weaker in Italy than in the US.

Data averaged by occupational level are utilised in Dell'Aringa and Lucifora [1994], who show how wage differentials by occupation started growing since the mid-1980s. The authors stress the importance of the wage drift for the reopening of differentials: a higher flexibility in pay determination on the employers' side often translated into individual wage premia for high level white collar workers.

Casavola et al. [1996] utilise occupational wages at the firm level from the Social Security Institute merged with data from the *Centrale dei Bilanci*[8] on firms' R&D expenses. The main aim of this study is a direct test of the skill biased technical change hypothesis as an explanation of growing inequality. The authors show how the growth of differentials during the central part of the 1980s has been concentrated in the wage distribution of non-manual workers, with a peak in 1988 followed by a slowing down in the next couple of years (the period studied is 1986-1990). A variance decomposition analysis within and between cells defined by occupation and

[8] Literally Central of Balance Sheets, this is a data set organised by the Banks' Association collecting balance sheet information from their affiliates.

degree of R&D intensity reveals how the growth of differentials has been faster in firms with higher levels of R&D intensity, both between occupations and within white collar workers; however, no effects can be detected in the blue collar wage distribution, while one of the predictions of the hypothesis is that inequality should grow within groups holding fixed observable characteristics. Regression analyses show that the effect of innovation has been stronger on relative quantities (white/blue collar workers ratio in the firm) than on relative wages, suggesting that the supply of skilled labour is rather elastic at the firm level.

The work of Manacorda [1997] takes into direct account the distributive effects of the Scala Mobile using micro-data from the Bank of Italy. At the aggregate level, it is shown that the degree of coverage against inflation granted by the indexation system has been dropping over the 1980s. Using quantile regression techniques with fixed cohort effects, the author assesses wage dynamics at different points of the wage distribution (first, fifth and ninth decile), showing that while the wage drift has protected purchasing power at the top of the wage distribution, the reduction of indexation coverage has determined a deterioration of relative wages at the bottom.

Lucifora [1998] examines the dynamics of low-pay incidence, with the low-pay threshold defined as two thirds the median of the distribution for full-time dependent workers, using micro-data from the Bank of Italy and INPS. At the aggregate level, it is shown how the probability of holding a job below the indicated threshold tends to drop over the first half of the 1980s, rising afterwards, with different timings depending upon the data set utilised, mirroring the dynamics for the whole distribution documented in the literature reviewed above. Individual low-pay probabilities are then modelled using probit equations, where the discrete outcome of interest is the occurrence of a wage below the low-pay threshold. Results reproduce what we should expect from standard wage equations, with factors such as labour market experience, education, jobs in large firms and in the north of the country which significatively reduce the likelihood of low pay.

The work of Brandolini and Sestito [1994] takes into account family incomes from Bank of Italy micro-data. The authors highlight three phases in the dynamics of the income distribution during the period analysed (1977-1991): a first phase of compression (1977-1982), followed by a phase of re-opening (1983-87) and a new phase of slight compression (1989-1991). Syncronicity of these phases with the business cycle suggests the existence of pro-cyclicality in the evolution of income inequality.

1.4 Individual wage profiles and long term inequality: the relevance of mobility

The empirical evidence on wage inequality reviewed above implicitly refers to an aggregate perception of wage differentials and doesn't account for the fact that the dynamics of cross-sectional wage distributions are generated by the evolution of individual wage profiles within the intertemporal distribution of wages. Using a rather old metaphor (Atkinson and Cowell [1983]), it can be said that analyses of cross-sectional inequality provide a series of snapshots of the wage distribution over time, but do not capture its dynamics as could be done by a movie. Analyses of wage dispersion are focused on the intensity of wage differentials between individuals, but don't assess their duration (Blundell and Preston [1998]); they are informative about the proportions of the labour force which, at a given point in time, occupy pre-defined classes of the cross-sectional distribution, but don't tell us anything about the degree of individuals' mobility within such classes through time, i.e. about the probability that the hierarchical ordering defined by the wage distribution changes over time. By observing that a proportion p of the labour force earns wages below a given quantile of the cross-sectional distribution it can't be determined if the whole labour force has a probability p of experiencing low-pay during the working career, if a proportion p of the labour force is trapped below that quantile for the whole career or to what extent the actual situation is a combination of these two extremes (Stewart and Swaffield [1999]). To throw some light on these aspects it is necessary to analyse individuals' mobility within wage hierarchies.

The passage from the analysis of wage dispersion to the analysis of wage persistence and mobility implies a change of focus from the moments of the marginal (cross-sectional) wage distribution to the ones of the joint (intertemporal) distribution. It is then clear that it is necessary to observe the wage of the same worker over several points in time, which is feasible only if panel data are available. If one wishes to study the wage distribution at time t and t+k, the dynamics of the variance can be assessed from the cross-sectional waves for the two years, but a subset of individuals observed in both years is necessary to estimate the correlation coefficient, thus taking into account wage persistence.

The use of panel data and mobility measures is such that the relevant time horizon for judgements on the actual degree of inequality is given by the working career or portions of it. This is evident if one wants to define a measure of multi-period wage inequality. Let w_{it} be a wage measure observable for N individuals in T time periods and let $w_{i.} = (1/T)\sum_t w_{it}$ be a measure of long-term wages (assuming a unit discount factor). Then, $var(w_{i.}) = (1/T^2)(\sum_t \sigma_t^2 + 2\sum_t \sum_{s>t} \rho_{ts}\sigma_t\sigma_s)$, where σ_t^2 indicates the variance in year t while ρ_{ts} is the correlation coefficient between years t and s. Long-term inequality thus receives contributions both from point-in-time inequality and from the intertemporal correlation of wages (immobility); the larger the latter, the larger long term inequality. Moreover, even if point-in-time dispersion remains constant, a reduction of wage mobility would imply an increase in long-term dispersion which would not be picked-up by point-in-time measures.

Cross-sectional inequality and mobility are the two components of long-term inequality; the first assesses the placement of individuals across the classes of the marginal distribution at a given point in time, while the second has to do with variations of such placement over time. The two aspects are independent. Hart [1983] shows how a constant value of the variance over time is compatible with any value of the correlation coefficient. Moffitt and Gottschalk [1993] demonstrate that the probability of changes in hierarchical orderings is only a function of the correlation coefficient of wages, but doesn't depend on the variances.

The relationship between measures of multi-period income inequality and point in time dispersion is formally analysed by Shorrocks [1978a], where it is demonstrated how (assuming convex inequality indices) the value of inequality measures based on long-term incomes can never exceed the average of such measures computed on cross-sectional incomes, the equality between the two holding when individual positions relative to the mean income do not change over time. This result follows intuitively from the fact that by raising the time interval over which long-term incomes are computed, the probability that individuals modify their hierarchical ordering becomes larger. Only in the absence of income mobility, i.e. when relative positions are fixed through time, point-in-time dispersion is an adequate indicator of long-term inequality. Shorrocks proposes an index of distributive rigidity based on the ratio between inequality of long-term incomes and the weighted average of point-in-time inequality measures.[9]

[9] See Björklund [1993], Arkes [1998] and Jarvis and Jenkins [1998] for recent applications of the Shorrocks measure.

By analysing the extent to which changing wage inequality is accompanied by variations in wage persistence, analyses of mobility can help in understanding the causes of rising inequality (see, for example, Moffitt and Gottschalk [1993]). Increasing wage persistence implies that inequality has to do with the remuneration of permanent workers' characteristics such as skills, so that theories of wage inequality based on variations in the relative demand for skills could be adequate in explaining the widening of the wage distribution. On the other hand, if rising inequality is parallel to increasing mobility, widening cross-sectional differentials wash out (at the individual level) after few periods, signalling a growth in wage instability which could arise from increased instability in the labour market, as put forward by theories emphasising the decline of labour market institution.

The dynamic perspective offered by studies of wage mobility also makes them relevant to assess the welfare consequences and policy implications of cross-sectional inequality. As an example, one of the policy measures suggested to cope with the increased incidence of low-pay jobs induced by rising wage inequality is the introduction of minimum wages to protect low labour incomes. Opponents of such a policy reply that low-paid jobs act as an entry point into the labour market, which is then abandoned thanks to wage dynamics generated by the acquisition of skills and experience (Bingley and Westergård-Nielsen [1995]), and maintain that minimum wages would increase rigidity and unemployment in a weak segment of the labour market. The validity of this line of argument depends on the extent with which wages are mobile over time: only if wages are mobile those observed in the lowest quantiles of the distribution will change status over time, so that the labour market is capable of dynamically redistributing incomes. It follows that the adequate policy in this case is the social diffusion of the workers' attributes generating mobility, rather than minimum wages.

An alternative way to see the relevance of wage persistence and mobility for policy purposes is suggested by Blundell and Preston [1998]. They observe that what matters for individual welfare is the level of consumption which, assuming perfect capital markets, depends on the expectation of long-term income (permanent income). As we saw above, long-run income inequality depends both on point in time inequality and intertemporal persistence. Then, an increase of cross-sectional inequality will affect permanent income and generate welfare losses the more it is paralleled by increases in correlation and persistence of the intertemporal income distribution; on the other hand, temporary fluctuations will only have a limited impact.

The link between low-paid jobs and mobility is relevant also for incentives to invest in human capital. The presence of wage dispersion is usually deemed a factor which stimulates the acquisition of skills (see, among others, Blau and Kahn [1996]). Hart [1983] argues that such an effect is not to be ascribed to the presence of differentials remunerating skills, but rather to the individual perception of the actual probability of reaching high wage positions, which in turns depends on the degree of mobility.

The above examples suggest that by reducing the impact of cross-sectional dispersion on life-time incomes, mobility should be desirable from a welfare viewpoint. However, wage mobility could also have negative effects. In particular, by increasing the volatility of individual wage profiles, it makes the outcome of individual efforts more uncertain, thus depressing incentives to invest in human capital (Hart [1983]; Guillotin and Bigard [1992]; Jarvis and Jenkins [1998]). Also, in the presence of risk aversion, wage volatility could induce intertemporal substitution in consumption, thus reducing households' welfare (Blundell and Preston [1998]). Hence, judgements on the desirability of wage mobility are not clear-cut. As suggested by Bingley and Westergård-Nielsen [1995], some insights into this trade-off may be gained by investigating the extent to which wage mobility is determined by permanent workers' characteristics: the idea is that while transitory wage fluctuations increase the uncertainty about the development of wage profiles, permanent movements are more predictable thus reducing negative effects.

Contrary to the analysis of wage dispersion, international comparisons of mobility are still rather rare. According to the 1996 OECD Employment Outlook, where some European Countries and the US are analysed, the degree of wage mobility tends to be pronounced and homogeneous across countries, with differences which do not mirror the ones in wage inequality. Such an outcome is confirmed in the subsequent issue of the same report (OECD [1997]), where it is also noticed that some sub-groups of workers (for example the more senior) tend to persist in low-paid jobs, once such a status has been experienced.

The remaining part of this Chapter is devoted to the empirical literature on wage persistence and mobility, distinguishing between the two research areas relevant to this Thesis: variance components models and the econometric modelling of wage transition matrices.

1.5 Variance components models and the distinction between permanent and transitory inequality

The use of variance components models of wages for the assessment of mobility can be introduced within the framework of classical panel data models with individual effects. Let us consider the "random effect" version of a panel log-wage equation, in which the error term is the sum of an individual specific component whose identification is possible thanks to the longitudinal structure of panel data, and by a white noise (WN) transitory component, independently and identically distributed (i.i.d.) with respect to both individuals and time periods, independent from the individual component; both components have a null mean with respect to individuals:

$$\begin{aligned} & w_{it} = \mu_i + \varepsilon_{it} \\ & \mu_i \sim (0, \sigma_\mu^2) \\ & \varepsilon_{it} \sim WN(0, \sigma_\varepsilon^2) \\ & E(\mu_i \varepsilon_{it}) = 0 \\ & i = 1, \dots, N \\ & t = 1, \dots, T \end{aligned} \tag{1.1}$$

where w_{it} is the residual log-wage, μ_i the permanent component and ε_{it} the transitory one. Residual variance is then composed by the sum of the variances of permanent and transitory effects; this last component vanishes when intertemporal wage covariances are taken into account:

$$E(w_{it} w_{is}) = \begin{cases} \sigma_\mu^2 + \sigma_\varepsilon^2 & \text{if} \quad t = s \\ \sigma_\mu^2 & \text{otherwise} \end{cases} \tag{1.2}$$

Given this framework, the intertemporal correlation coefficient is constant and equal to the share of permanent on total variance; recalling that the correlation coefficient is a measure of immobility, it is clear how this class of models can be used for the analysis of wage mobility.

Figure 1.1 gives a representation of wage dynamics resulting from (1.1) and (1.2) for two hypothetical individuals (A and B) by different degrees of permanent wage dispersion (cases 1 and 2). Straight lines parallel to the time axis represent the levels of permanent wages, which is to say, the permanent components constant over time; in this model, residual permanent wages are characterised by individual specific intercepts, while growth rates are assumed to be zero for both individuals.

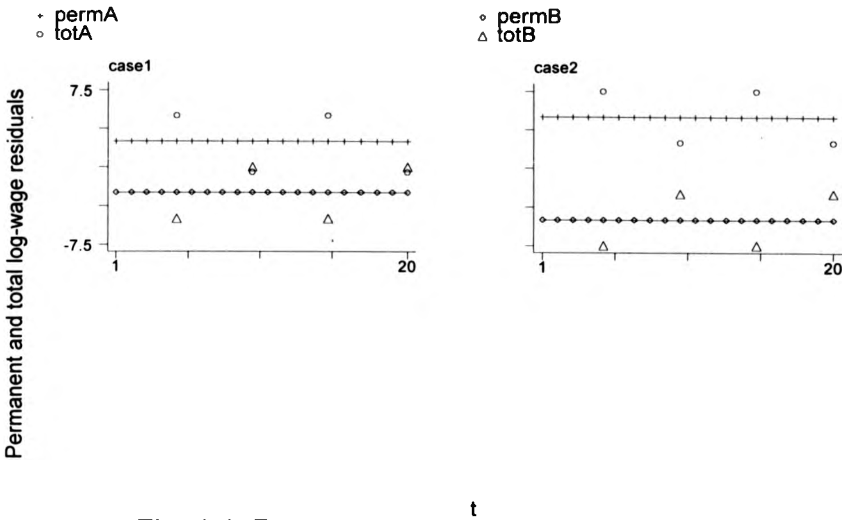

Fig. 1.1: Permanent wage inequality and mobility

The dots scattered across permanent wage lines represent total wages; in each of the four time period considered, the distance between the dot and the straight line gives the transitory wage fluctuation. The left panel (case 1) shows how individual's B permanent wage is lower than individual A's one; however, transitory fluctuations are such that in some periods, her total wage equates the one of individual A. By comparing total wages in the first and the last period observed, a change in hierarchical ordering takes place, generating mobility. The right panel maintains the same degree of transitory wage volatility but the distance between the two permanent wage profiles is now higher, so that case 2 is characterised by a larger permanent variance. In this case, transitory shocks are not capable of modifying the wage rank: an increase in permanent inequality generates, *ceteris paribus*, a reduction of mobility.

Lillard and Willis [1978] apply the model above to PSID data, allowing the transitory component to follow an AR(1) process, showing, among other things, how incomes for the black sub-sample present higher persistence than the white one. Their study is focused on the analysis of low-pay probabilities; making some additional hypotheses on the distribution of permanent and transitory components (in particular, they assume normality for both), these authors decompose the probability of the low-pay status into permanent and transitory components, showing, again, that they are more persistent for the black sub-sample.

A series of contributions to the literature on variance components models is characterised by a more flexible specification of the permanent component, which is specified as a linear function of time or labour market experience, so that each individual is described both by specific intercept and slope parameters and heterogeneity is not only due to fixed unobserved factors, but also to differences affecting the dynamics of wage profiles, such as heterogeneous learning abilities.[10] Such a specification is known as random growth model (Moffitt and Gottschalk [1993]; Baker [1997]), meaning that growth rates are a random variable distributed over individuals. A central role in this model is played by the covariance between intercepts and slopes of individual wage profiles. A negative value of this parameter implies that those observed in the bottom part of the initial wage distribution will experience larger growth rates than those at the top and *viceversa*, so that the distribution of permanent wages will tend to converge over time. Such a prediction

[10] It is worth stressing that since such a specification refers to residual wages, usually net of quadratic trends in age or labour market experience, assuming linearity doesn't contradict the prediction of wage profiles concave over the life-cycle arising from human capital theories.

arises, for example, from the Mincerian theory of (generic) on-the-job training. Those who invest in the acquisition of skills employable also outside the firm where the training takes place bear, at least in part, the cost of the investment and receive initial wages which are lower compared to those not investing. Since the present discounted value of wage flows must be equal across alternatives, the wage profile of those investing must be steeper, so that low intercepts will be associated to high rates of growth. This literature has based tests of the Mincerian hypothesis on the sign of the covariance parameter. From the mobility view point, a negative covariance implies that permanent wage profiles will cross each other over the working career, generating mobility and an overall equalisation of permanent wages. Hence, the random growth model provides insights on the extent of permanent wage mobility, which we have seen in Section 1.4 to be relevant for judgements on the welfare impact of overall mobility.[11] It has to be stressed that other theories can generate predictions on the sign of the covariance parameter. As an example, combining signalling and matching models, initial wages will depend upon workers' attributes observable by the employer at the beginning of the match, while wage profiles will be influenced by the revelation of information concerning the quality of the match over time. Assuming that workers who have better observable characteristics at the beginning are indeed more productive, those with high initial wages will also experience high growth rates: the covariance between intercepts and slopes will be positive in such a case, so that the distribution of permanent wages will diverge over the life-cycle.

Lillard and Weiss [1979] consider a sample of American scientists and find a negative covariance in raw wages, with the parameter which turns out to be positive

[11] In the simplest error components model of wages in (1.1) mobility depends on the size of transitory variance relative to the total one, but no room is made for mobility within the permanent wage component.

when age and time effects are removed, implying that permanent inequality grows over the life-cycle within the cells defined by these characteristics. These authors also present results by birth-cohort, showing that the dispersion of starting permanent wages is higher for older cohorts. Hause [1980] bases a test of the Mincerian on-the-job training hypothesis on a sample of young Swedish males; his results indicate a negative value of the covariance between intercepts and slopes, supporting the hypothesis. Bourguignon and Morrisson [1983] analyse a sample of French white collar workers and utilise a quadratic specification of the permanent wage, finding that the covariance is positive between intercepts and second order terms, negative between intercepts and first order terms and negative between first and second order terms. The sign of these covariances are in accordance with the presence of cross-overs of wage profiles, and the authors stress that this implies mobility of permanent wages. A renewed interest in this class of models has emerged in recent years. Baker [1997] uses PSID data and compares the Mincerian model with a random walk specification of the permanent wage, providing evidence supporting the former; his estimates imply that individuals with a slope parameter one standard deviation above the mean will experience a 20-30% growth of permanent wages over the analysed 10 year period. Baker and Solon [1998] utilise Canadian panel data and introduce also a quadratic specification of the transitory variance with respect to age, aimed at capturing a larger wage instability at the beginning and the end of the working career. Results show both permanent wage convergence and a convex profile of transitory variance over time.

An element of differentiation within the literature focused on random growth models is given by the estimation technique; in particular, while studies produced between the 1970s and the 1980s utilise GLS or maximum likelihood estimators, more recent contributions are based on the minimum distance estimator

(Chamberlain [1984]). This is, in practice, an application of the GMM with which the structure of second moments implied by the theoretical model is mapped directly into empirical second moments, according to some metric. [12] One of the first applications of this method to the longitudinal analysis of wage dynamics is given in Abowd and Card [1989]. These authors are interested in the covariance between wages and hours of work and use the PSID to estimate the relevant parameters from first order differences. They show that the covariance between changes in wages and changes in hours is positive within the same period, becoming negative for data one year apart and negligible afterwards.

The application of the minimum distance estimator is a qualifying aspect of two recent contributions to the literature on the wage covariance structure (Moffitt and Gottschalk [1993] and Diçkens [1996]) which focus on the permanent vs transitory nature of the rise in wage dispersion characterising the UK and the US in recent years. Both studies put great emphasis on the modelling of shifts in the covariance structure through time, which is pursued by introducing a time varying loading factor on each wage component, thus introducing time-wise heteroskedasticity.[13] The dynamic analysis of the two wage components allows an assessment of their role within the aggregate dynamics of wage dispersion. From a policy perspective, this provides an estimate of the share of total inequality growth accounted for by the permanent component, thus indicating to what extent rising inequality persistently affects individual wage careers. As stressed above, this is also relevant for interpretation purposes: while increases in the relative importance of the permanent component support skill based explanations of changing wage inequality, a growth in

[12] The method will be discussed in detail in the next Chapter. The method is not confined to the analysis of second order moments but allows a joint analysis of first and second order moments (see Crepòn and Mairesse [1997]).

[13] As pointed out by MaCurdy [1982], this doesn't require any explicit parametrisation aimed at dealing with the incidental parameters problem, since, in the context of panel micro-data, period specific parameters are identified by cross-sectional variability.

volatility could result from an increased instability of working careers, induced, for example, by a lower degree of institutional wage protections.[14]

Moffitt and Gottschalk [1993] use PSID data and adopt ARMA(1,1) specifications of the transitory wage, which allow wage rigidity to depend not only on the relative importance of the permanent wage, but also to shocks persistence. These authors show that the growth of US wage differentials has predominantly affected the permanent component during the 1970s; the 1980s are instead characterised by a different scenario, wage volatility being more relevant. Their results suggest an increasing instability and competition in the labour market in recent years. Similar conclusions are reached by Dickens [1996] who uses NES data and shows that growing UK wage inequality mainly arose from permanent differentials from the mid 1970s to the mid 1980s, while, afterwards, wage volatility has become increasingly important.

1.6 Transition matrices, mobility indices and econometric models of transition probabilities

The methodological approach to the analysis of wage mobility based on transition matrices directly focuses on the probability that the hierarchical ordering defined by the cross-sectional wage distribution changes through time. The starting point in this case is the discretization into classes of the marginal wage distribution of the two years delimiting the transition analysed. The availability of panel data allows

[14] As mentioned earlier in this Chapter, the distinction between permanent and transitory sources of the rise in US wage inequality is also at the core of the analysis in Gottschalk and Moffitt [1994], who do not utilise formal models of the covariance structure, but identify the permanent wage as a multiperiod mean of yearly wages and compute the transitory wage as difference between such means and yearly wages. They show that about one third of the rise in inequality over the 1970s and the 1980s can be ascribed to the transitory component. In his comment to this paper, Katz [1994] points out that, due to the introduction of technical innovations, workers in their old job can behave as new workers, presenting a higher degree of wage instability, a *caveat* which should be borne in mind when deriving interpretations on the factors causing growing wage differentials from these models.

crossing of these two orderings, yielding a contingency table whose elements represent the absolute frequencies of the intertemporal (discrete) wage distribution. Standardisation of each cell of the matrix with respect to the absolute frequencies of the marginal distribution of the starting year yields a transition matrix, which describes the probability of belonging to the classes of the arrival wage distribution conditional on starting classes.

1.6.1 Mobility indices

The classic way by which transition matrices have been employed to analyse wage mobility is based on mobility indices (see, for example, Atkinson et al. [1992]). A formal discussion of the theory of mobility indices is provided by Shorrocks [1978b], where mobility indices are defined as functions mapping from the space of transition matrices to the real axis, in particular to the unit interval in case of normalised indices. Some desirable properties for these indices are defined and the concepts of *complete immobility* and *perfect mobility* introduced. The first case corresponds to the absence of hierarchical permutations, with the resulting transition matrix given by the identity matrix. In the second case, wages of the arrival year are stochastically independent from those in the starting year, so that transition probabilities are independent from the starting class and the rows of the transition matrix are all equal.[15] The two extreme cases allow standardisation of indices computed under different circumstances, so that values of the same index computed on distributions with differing number of classes[16] or values of different indices can be compared.

[15] The definition of perfect mobility usually adopted for operational purposes is more demanding and also requires that transition probabilities are equal across arrival classes (what Shorrocks [1978b] calls *strong perfect mobility*). Assuming that the marginal distribution is constituted by p classes, the resulting transition matrix has all elements equal to 1/p.

[16] The choice of the classes to adopt in the discretization of the marginal distribution is an obvious source of arbitrariness (Atkinson et al. [1992]). Alternatives commonly adopted are equal partition of the wage range, classes defined as ratios on the mean or median of the distribution or quantiles. The last

Commonly used mobility indices[17] can be classified according to the aspect of mobility analysed, namely the frequency or width of transitions. A basic frequency indicator is the *immobility ratio*, equal to the ratio between the trace of the transition matrix and the number of wage classes. Given that frequencies on the diagonal correspond to individuals who do not change wage rank during the transition, this measure gives the average probability of persistence. A more demanding definition of mobility is behind the *quasi-immobility ratio*, which considers also elements on the two main sub-diagonals; for an individual to be classified as mobile it is then necessary to cross the borders of at least two wage classes. Further specialisations of frequency indicators take into account the direction of movements; this is the feature of the *ascending* and *descending mobility ratios*, which are based on frequencies in the super- and sub-diagonal part of the matrix, respectively. The *expected absolute jump*, i.e. the average width of transitions conditional on having moved, is the most frequently adopted width indicator. Similarly to frequency indicators, also in this case the index can be specialised to take the direction of transitions into account.

The immobility ratio computed by starting class is used by Shorrocks [1976] to test the hypothesis that personal incomes evolve according to the first order Markov hypothesis. The implication of the theory under scrutiny is that the matrix governing the two-periods transition is the product of the two one-period matrices.[18] By

option corresponds to a relative view of the mobility process. Quantiles usually have varying width and are typically more compressed towards the middle of the distribution, as long as frequencies tend to be concentrated over the central part of the wage range. They present the advantage of being robust to the presence of outliers and of being equally representative across the classes of the starting wage distribution.

[17] See Bourguignon and Morrisson [1983] and Guillotin and Bigard [1992] for an overview on mobility indices and their formalization.

[18] The Markov hypothesis states that the level of current incomes and the transition matrix are sufficient for the prediction of future incomes, without additional hypotheses on the past of the income process; also, state and transition probabilities are assumed to be independent. Let us consider three points in time (i,j,=1,2,3) and let p_i be the row vectors of the marginal wage distributions, and Q_{ij} the corresponding transition matrices. The hypothesis implies that: $p_2=p_1Q_{12}$; $p_3=p_2Q_{23}$; $p_3=p_1Q_{13}$. It follows that $Q_{13}=Q_{12}Q_{23}$.

comparing immobility rates from the two-periods matrix with the ones from the product of single step transitions, the author shows that in this last case income mobility appears to be higher, rejecting the first order Markov hypothesis. Bourguignon and Morrisson [1983] use both frequency and width indicators on their panel of French white collars. They show that by raising the time span over which such measures are computed, mobility increases at decreasing rates. French data are analysed also by Guillotin and Bigard [1992], where attention is focused on the disaggregation of mobility indices by starting and arrival wage class and a methodology for the identification of immobility poles within the whole distribution is proposed. Their results show a considerable degree of mobility over the 15 years period analysed, which tends to drop as higher ventiles are taken into account. A comparison between French and Italian data using similar methodologies is provided by Bigard et al. [1998], showing that the Italian distribution is more rigid, especially at the bottom. Dickens [1997] utilises immobility ratios on NES data, showing that the increasing wage inequality has been accompanied by drops in mobility. The author also proposes a mobility measure based on changes in percentiles for each workers, corresponding to a mobility matrix in which each individual has her own wage class. Finally, immobility ratios are computed in OECD [1996] (whose results have been commented above) to produce international comparisons of mobility.

1.6.2 Econometric modelling of transition probabilities

The literature on transition matrices concentrates on the statistical description of mobility. Recent years have instead been characterised by studies focusing on the econometric modelling of individual transitions within the classes of the wage distribution; in other words, these contributions take into account the determinants of the assignment of individuals to the cells of the transition matrix. Attempts in this direction can also be found in the descriptive literature, where they are based on the estimation of mobility indices on sub-samples defined according to workers' attributes (see, for example, Bigard et al. [1998]). Such an approach presents at least two disadvantages. Firstly, the analysis remains at the aggregate level in that it cannot distinguish between heterogeneity and pure state dependence within aggregate persistence, an important distinction which will be discussed at length in Chapter 4. Secondly, the interpretation of results is not straightforward, in particular being hard to make *ceteris paribus* statements in the presence of many explanatory variables.

Given that the object of the analysis is a probability, the natural route followed by researchers for its econometric modelling is the one of discrete response models. In practice, it is supposed that for each individual the discrete outcome mobile/not-mobile is the realisation of some latent propensity to move, of which only a binary realisation is observable; the resulting dummy indicator is then regressed on the set of personal characteristics by means of discrete response models. Compared to an equation of wage growth rates, such an approach introduces a loss of information by discretising an originally continuous variable, so that no distinction is made between transitions of different length. However, this caveat can be overcome by using discrete response models with more than two feasible outcomes to assess the number of jumps such as ordered probits or count data (see Chapter 4 for an application of the first).

The econometric analysis of transition probabilities is complicated by two potential sources of endogenous selection which are inherent to the structure of the problem. First of all, analysing transitions requires conditioning on lagged wage states and, as long as the assignment of workers to such states is non-random, selecting those starting from a given wage class to estimate the model can bias parameters estimates. Secondly, transitions can be observed only for those who belong to the sample of wage earners at both extremes of the time interval investigated, and the presence of non-random exits from the wage distribution can, again, lead to biased estimates. A first group of studies assumes the exogeneity of both selection processes. This is the case in Smith and Vavrichek [1992], who adopt a logit model to analyse mobility out of US minimum wages for a sample of individuals previously employed at the minimum wage and show that the lack of education is the main factor in determining the ceteris paribus probability of persistence. A linear probability model, in which the dummy dependent variables is treated by OLS, is proposed by Gregory and Elias [1994] to analyse transition probabilities from the bottom quintile of the distribution using NES data: they find that low-pay persistence rises with age. Sloane and Theodossiou [1996] use BHPS data and adopt a multinomial logit model to analyse destination states of workers observed below the third decile of the origin wage distribution, with destinations also including the exit from the data set, in this way avoiding selection on the basis of panel retention. It is shown that low-wage persistence is significantly lower for males, workers in large firms or those participating in re-training programmes. A logit model is utilised by Contini et al. [1998] to model the conditional probability of being below the third decile of the distribution for those starting the transition below the third decile or above the seventh on Italian administrative data (INPS); they show that job mobility and employment in large firms positively influence transitions out of low-pay, while, on the other hand, employment in the service sector and past unemployment episodes favour transitions into low-pay. Guillotin and Hamouche [1998] utilise a count data model to analyse the number of ascending jumps through deciles computed on French administrative data (DADS). They consider the whole origin distribution and include dummies for the origin decile among regressors, thus conditioning on starting states. Their results suggest that human capital accumulation favours ascending mobility and preserve higher wage ranks once reached.

The formal assessment of the potential endogeneity of initial conditions of the wage process and attrition from the wage distribution characterises two recent contributions to the econometric literature on wage mobility. The set-up adopted in these studies is based on multivariate microeconometric models in which selection and transition probabilities are jointly estimated, thus allowing correlation between error terms and tackling endogeneity problems. The framework is basically that of endogenous selection models, in which control is made for the fact that the equation of interest (the mobility equation) can be estimated only if observations satisfy two endogenous sample selection rules (initial conditions and attrition). The work of Bingley et al. [1995] jointly models state probabilities, retention probabilities and mobility across deciles of the wage distribution using a panel of Danish wages: the whole set-up is a trivariate probit. Identification of the model requires exclusion restrictions in the form of variables only entering each of the selection equations and not the mobility equation or the other selection equation; the authors assume that age only affects retention probabilities, while industrial affiliation and the number of children in the household only affect the origin decile. Results from the mobility equation indicate that changes in occupation or industry determine downward mobility; on the other hand, human capital in the form of education and labour market experience is associated with upward movements. The estimated error covariance matrix is characterised by statistically significant elements, signalling the actual endogeneity of the two selectivity processes and justifying the whole framework adopted.

The paper of Stewart and Swaffield [1999] is focused on the analysis of low-pay transition probabilities using BHPS data. They use a bivariate probit to model the probability of current low-pay states conditional on lagged states, thus tackling the endogeneity of initial conditions. Their base model considers only wage earners at both extremes of the transition; the control for panel attrition is assessed by extending this model and modifying the binary outcome in the arrival year from "low-pay/high-pay" to "not moved up/high-pay", where "not moved up" includes low-pay and exits from the wage distribution. The authors pay great attention to the strategy adopted for identification of initial conditions; in particular they reach exact identification using the square of labour market experience (which doesn't enter the transition equation given its nature of wage change equation) and test the validity of parental background indicators as instruments for initial conditions, concluding in favour of their use. This paper is also characterised by the parallel use of several low-pay thresholds, and it is shown that results are typically robust to changes in the definition of low-pay. The estimated correlation coefficient between state and transition probabilities is statistically significant, thus rejecting the exogeneity of initial conditions; by comparing results from the bivariate probit with those obtained assuming exogenous initial conditions the authors show that in this last case both size and significance of estimates is higher, especially for education. On the other hand, factors such as training, plant size, union coverage and gender retain their significance in affecting low-pay transition after controlling for endogenous initial conditions.

Chapter 2

The covariance structure of male wages

2.1 Introduction

This Chapter applies the minimum distance technique to analyse the covariance structure of an unbalanced panel of Italian male wages. As stressed in Chapter 1, the analysis of the intertemporal covariance structure allows decomposition of cross-sectional wage inequality into a part due to permanent workers heterogeneity and a part due to the volatility of wage shocks, a distinction relevant both for interpretation purposes and to draw policy indication. From the first viewpoint, while permanent wage differentials are usually associated with the remuneration of permanent skills, wage volatility is more in line with a growth of labour market instability, which could in turn result from the decline of labour market institutions. Thus, analysing which of the two components is predominant in determining aggregate cross-sectional inequality can shed light on the factors driving changes in wage differentials. For what concerns the policy implications, permanent inequality implies that cross-sectional wage differentials are likely to persist over the life-cycle, making the need for low-wage protection more urgent.

The permanent wage will be specified in this Chapter according to the random growth model, in which individual (residual) wage profiles are assumed to be linear, and whose second moments allow assessment of the extent of cross-overs in permanent wages, and, consequently, of permanent wage mobility, thus providing insights on the welfare impact of overall mobility (see the discussion of Section 1.4). Also, time-wise heterogeneity of the covariance structure will be allowed by introducing time varying loading factors within each wage component.

The covariance structure analysis will be based on an unbalanced panel made available from the National Social Security Institute for the 1974-88 interval. The data thus cover a period of remarkable changes in wage determination policies for the Italian labour market (recall the discussion of Section 1.3), going from the era of wage egalitarianism to the second half of the 1980s, when pay determination on the employers' side became more flexible and the use of individual wage premia became more frequent, especially in favour of white collar workers.[19]

The Chapter is organised as follows. Section 2.2 shows how the random growth model of residual wages can be implemented within the minimum distance estimator framework. Section 2.3 describes the data utilised, while Section 2.4 reports the results obtained. Section 2.5 summarises the main findings of this Chapter, while details on the estimation method and the STATA codes written to implement it are outlined in the Appendix.

2.2 Model specification

The minimum distance technique will be applied in this chapter in order to model the second moments of the joint intertemporal wage distribution on an unbalanced panel of Italian wages made available by the National Social Security Institute (*Istituto Nazionale di Previdenza Sociale*, INPS). In particular, the distance to minimise will be the unweighted sum of squared deviations of the theoretical covariance structure from the empirical one (Equally Weighted Minimum Distance, EWMD). Estimated models of earnings residuals[20] will derive from the following general specification:

[19] Such a data set was the only one available from the Institute by the time this research started. The analysis of a new draw from the same archive, covering a more recent period, will be the object of Chapter 3.
[20] Before proceeding to the minimum distance estimation, raw wages are first adjusted for the effects of calendar time, age and birth cohort; details of these adjustment are given in section 2.4.

$$\begin{aligned}
w_{it} &= g(t)(\mu_i + \gamma_i t) + h(t)v_{it} \\
v_{it} &= \rho v_{it-1} + \varepsilon_{it} + \theta\varepsilon_{it-1} \\
\varepsilon_{it} &\sim WN(0,\sigma_\varepsilon^2) \\
\begin{pmatrix}\mu_i \\ \gamma_i\end{pmatrix} &\sim \left[\begin{pmatrix}0 \\ 0\end{pmatrix}, \begin{pmatrix}\sigma_\mu^2 & \sigma_{\mu\gamma} \\ & \sigma_\gamma^2\end{pmatrix}\right] \\
t &= 0,1,\ldots,14
\end{aligned} \tag{2.1}$$

where w_{it} is the log-wage residual for individual i in year t, the core permanent component is assumed to be linear in calendar time (μ_i+γ_it), the core transitory component (v_{it}) is specified as an ARMA(1,1) process[21] and both are loaded by time-varying loading factors, g(t) and h(t) respectively. Hence, each wage component is composed by a core, representing individual heterogeneity, and a loading factor, which inflate or deflate the distribution of the cores but doesn't alter relative rankings within it. For the purposes of the present Chapter, the loadings will be specified as cubic polynomials in calendar time.

The first parameter of the core permanent wage component (μ_i) represents omitted individual characteristics whose effect is fixed over the working career, such as environmental background, while the second (γ_i) represents unobserved heterogeneity which affects earnings growth rates, such as learning ability. As stated in Section 1.5, a central role in determining wage dynamics is played by the covariance of these two terms across individuals; a negative value of $\sigma_{\mu\gamma}$ implies that wage growth over the life-cycle operates as a compensatory mechanism off-setting time invariant components of permanent inequality, as individual positions within the permanent wage distribution at the beginning of the period observed will be inversely related to future positions. The resulting cross-overs of permanent wage profiles

[21] Moffitt and Gottschalk [1993] stress that in the case of autoregressive specifications of the transitory wage it would be more appropriate to refer to a mean-reverting wage component, rather than to a purely transitory one. Bearing this point in mind, in the analyses of the current and the next Chapter I will still refer, for ease of exposition, to transitory wages.

imply that mobility doesn't increase wage uncertainty, thus avoiding the welfare worsening effects mentioned in Section 1.4.

The theoretical covariance structure implied by the model is:

$$\begin{aligned} E(w_{it}w_{is}) = {} & g(t)g(s)[\sigma_{\mu}^2 + \sigma_{\gamma}^2(ts) + \sigma_{\mu\gamma}(t+s)] + \\ & h(t)h(s)\{\rho^2 E(v_{it-1}v_{is-1}) + \sigma_{\varepsilon}^2[d(1+\theta^2) + d_1\theta]\} \end{aligned} \tag{2.2}$$

$$\begin{aligned} d &= I(|t-s| = 0) \\ d_1 &= I(|t-s| = 1) \end{aligned}$$

where I(A) is a dummy indicator equal to 1 when A is true and 0 otherwise.

Some remarks arise. First, permanent inequality is determined by initial fixed heterogeneity (σ_{μ}^2), heterogeneous life-cycle behaviour (σ_{γ}^2) or from the covariance between the two ($\sigma_{\mu\gamma}$); the first two terms contribute positively to permanent inequality, while the third contribution could be negative. Moreover, earnings covariance between two points in time is lower the lower the value of $\sigma_{\mu\gamma}$. Thus, the presence of a negative covariance between individual intercepts and slopes implies a convergence process of permanent wage levels (permanent mobility) which lowers permanent inequality over individual life-cycles without raising volatility.

Secondly, it has to be noted how core permanent wages are specified as a function of calendar time rather than labour market experience, which is the natural variable one would have in mind in a human capital wage profiles framework. Nevertheless, neither actual labour market experience nor years of education were available in the INPS data (see section 2.3 below) and it was not possible to construct a measure of potential experience. However, as we will see below, wages are net of age and birth cohorts effects, so that the use of calendar time is not such a dramatic simplification.

Thirdly, it can be noted how the core permanent wage specification places the intercept at the first year of the sample period, thus assessing wage convergence by considering the covariance between initial and life-cycle heterogeneity. Other studies (in particular, Lillard and Weiss [1979]) consider calendar time in deviations from the sample mean ($\bar{t}$) so that fixed heterogeneity is measured at $t=\bar{t}$ rather than at $t=0$. Such an alternative specification yields a different interpretation of the covariance between fixed and life-cycle heterogeneity. To see this, let us rewrite our specification of the random growth model (denoted as w_{it}^{P}):

$$w_{it}^{P} = \mu_i + \gamma_i \bar{t} + \gamma_i (t - \bar{t}) \tag{2.3}$$

where now the new intercept $\mu_i^{*} = \mu_i + \gamma_i \bar{t}$ also includes a component representing the effect of heterogeneity in wage growth on residual wages at $\bar{t}$. Clearly: $E(\mu_i^{*}, \gamma_i) \equiv \sigma_{\mu^{*}\gamma} = \sigma_{\mu\gamma} + \bar{t}\sigma_{\gamma}^{2}$ and a negative $\sigma_{\mu\gamma}$ may well be consistent with a positive $\sigma_{\mu^{*}\gamma}$, the two specifications apparently yielding contradictory conclusions. In fact, such an occurrence would be observed for $\sigma_{\mu\gamma} \in (-\bar{t}\sigma_{\gamma}^{2}, 0)$, i.e. when initial convergence is "weak" enough that its effects vanish between 0 and $\bar{t}$. It can also be observed how in case of initial divergence, conclusions arising from the two specifications are always in accordance.

A final remark has to be made on the core transitory component of the autocovariance structure in (2.2). The presence of an autoregressive component implies that current moments depend on their lagged values and, if one traces the recursion back to the first year of the sample, the whole theoretical covariance structure will depend, in its transitory part, on the moments of the initial conditions of the stochastic process. The usual assumption in time series analysis is that such initial conditions are irrelevant, or, equivalently, that the process started in the infinite past. As pointed out by MaCurdy [1982], such an assumption is untenable in panel data analysis, where the number of time periods is relatively small, and could lead to biased parameter estimates. A way to tackle the problem in the GMM context (see also Baker [1997]) is to include an additional parameter representing the accumulation of the stochastic process up to the first year of the panel which explicitly models the distribution of initial conditions, denoted σ_0^2 in the tables reporting the estimation results.[22]

2.3 The data utilised

The data source is the INPS 01/M form, a form that each employer has to fill, for each of her employees, in order to pay the National Social Security contributions. The data set has not been originally constructed for research purposes, but comes as a by-product of the administrative activity. This implies a good reliability of the information collected, but, on the other hand, explanatory variables are relatively few[23].

The sample refers to full-time workers from the private non-agricultural non-self-employed sectors of the Italian economy over the 1974-1988 period. Each

[22] The way in which the covariance structure has been parameterised with respect to the initial conditions of autoregressive stochastic processes is described in the Appendix.

[23] The available information concerns, among others, the workers' sex, age and occupation and the firms' size, industry and geographical location. Education and labour market experience are not observed. For a full description of the explanatory variables available in the INPS data see Lucifora [1995].

worker is denoted by an INPS identification number which is constant over all the working career, enabling individual working histories to be traced over time, generating a panel. The available wage variable is the annual wage, comprehensive of extra-time payments and gross of income taxes; information on weeks worked during the year is also provided, and the analysis has been carried out in terms of weekly wages. For the purposes of this study, attention has been restricted to male workers aged between 23 and 65 and born between 1923 and 1951.[24] This yields an unbalanced panel where the total number of workers (i.e. entering the panel at least once) is around 18,000.

Potential attrition problems can affect the data. Reasons for movements into and out from the data set are:

- periods of unemployment;
- changes to and from self-employment status;
- retirement;
- mobility to or from the public sector (which, with few exceptions, is not covered by the INPS file).

Trying to find good instruments to model sample selection, whose likely effect is to understate wage volatility, is difficult due to the limited information available and no attempts have been made to solve this problem; however, the use of a data set with an unbalanced design partially mitigates it.

Table 2.1 describes the structure of the INPS panel. The diagonal shows how the cross-sectional size of the data set tends to increase from 1974 to 1985, while a reduction occurs in the last three years. The attrition path of each cross-section

[24] The upper limit of the birth cohort range comes from the original data set; the lower limit, which excludes those individuals ageing over 65 during the sample period, selects a negligible proportion of wage profiles (0.45%) out from the original data set.

through subsequent years of the panel is reported above the diagonal, where the number of observations used in the computation of each cell of the covariance matrix for the whole sample is collected: as can be seen, this path is, with few exceptions, decreasing.

The evolution of some sample features over the period is reported in Table 2.2. As can be seen, apart from a tendency to white collar undersampling in 1976 and 1977, the occupational structure tends to stay constant from 1974 to 1988, while the cohort composition shows how the weight of younger cohorts grows through years. The last two rows of the table report the evolution of the cross-sectional distribution of logarithmic wages over the period: mean logarithmic wages recorded a 5% growth over the period (23% in levels) while wage dispersion dropped of almost 20% from 1974 to 1982 and grew thereafter, the value in 1988 being approximately equal to that in the starting year. We turn now to the longitudinal analysis of such inequality dynamics.

2.4 The covariance structure of Italian male wages

This section contains the results of the empirical analysis of Italian male wages. The analysis has been carried out for the whole sample (blue collars, white collars and managers) and for blue and white collar workers separately; this last exercise seems particularly relevant, given that, as seen in the Section 1.3, the use of flexible pay determination schemes during the 1980s developed mainly for non-manual workers.

In order to construct the wages covariance matrix, real wages[25] have first been adjusted for year, age and cohort effects; the three effects are intended to capture

[25] Nominal figures have been deflated with the CPI (1985=100).

business cycle, life-cycle and productivity growth effects respectively and to identify the last two separately, cross-sections have been pooled over years. Age effects are controlled for by a quadratic function of age, while year and cohort effects are specified as dummy variables. Figure 2.1 plots kernel density estimates of the cross-sectional distribution of residuals from this regression for the whole sample: a noticeable feature is the concentration of frequencies around the mean taking place from 1977 to 1984, which results in an increase of the corresponding density height; this tendency reverts thereafter, especially during the last two years of the sample, when the peak's height decreases.

2.4.1 The empirical covariance structure

For each of the 3 groups considered (i.e. whole sample, blue collars and white collars[26]) residuals from such initial regressions have then been utilised to construct the respective covariance matrices; Table 2.3 reports the one for the whole sample together with each element's standard error, while the corresponding correlation coefficients are reported below the diagonal.

As can be noted in the table, the covariance structure fails to asymptote to a long run level, a feature which has been observed both in UK and US data (see Dickens [1996] and Moffitt and Gottschalk [1993] respectively): a marked historic-time dependence seems to be present in the data, with covariances peaking in 1985 and especially in 1988.

Figures 2.2 to 2.4 provide some further element of discussion by reporting the evolution of the variance (i.e. the diagonal of the covariance matrix) and of some

[26] The blue collar/white collar split has been carried out by considering the worker's occupational status of each year. This means that if a worker is classified as blue collar in year t and as white collar in year s (i.e. if a change in occupation takes place between t and s) he will not contribute to $E(u_{it}\ u_{is})$ for either of the two sub-samples.

wage persistence and immobility indicators, namely the correlation coefficient and the ventile quasi-immobility ratio.

The first panel of Figure 2.2 displays the well-known picture of trends in Italian wage inequality (see, for example, Dell'Aringa and Lucifora [1994]), with a strong compression of differentials during the 1977-82 period, in which the egalitarian system of indexation was fully effective, and a reopening thereafter, particularly marked in 1988. The remaining two panels of Figure 2.2 help in assessing the evolution of inequality within each occupational group; the reopening of wage differentials affects white collars data since 1984, while the decrease of blue collars' wages variance continues, although at a diminished pace, until 1987. If compared with the evidence in the top left corner of the figure, this suggests that the widening of overall wage differentials was, to a large extent, driven by inequality within white collar workers, for which it seems that the relaxation of equalising forces had quicker effects.

Patterns of wage persistence are described in Figure 2.3 by means of the correlation coefficient, which is plotted for wages one and five years apart. The graph in the top left corner shows a short term measure which tends to cycle before 1982 and to slightly decrease in the years of growing dispersion, with the exception of the last year of the sample; the picture is different for the medium term measure, for which a marked drop can be observed from 1985 onwards. Taking manual wages into account, it can be seen how correlation tends to be concentrated in the central-final part of the sample period, depending upon the lag width considered; both the decrease in short term correlation since the early 1980s and the drastic drop in medium term correlation from 1985 onwards are still evident. The data for white collar workers data tell a somewhat different story, with correlation which increases both in the short and the medium term during the last years of the data, with a drop in 1988.

Figure 2.4 plots the values of the ventile quasi-immobility ratio for wages one and five years apart; compared with an usual immobility ratio (i.e. the average, across classes, proportion of cases not changing wage class during the transition considered) this measure considers immobile also those individuals moving only to the adjacent class and is thus robust to the effect of small wage "pushes".[27] Compared to the correlation coefficient, which is based on co-deviations from the marginal mean and picks-up absolute changes at each point of the wage range, quantile mobility indices measure variations in relative ranks and ignore within classes movements (see also Jarvis and Jenkins [1998]).

The graph shows, for the whole sample and blue collar workers, a tendency for the frequency of transitions among ventiles to initially drop and then level-off towards the central years of the data, especially in the short term. The pattern is slightly different for the white collars' sample, where immobility steps up in correspondence of the increase in autocorrelation observed above.

2.4.2 Variance components models for the covariance structure

Table 2.4[28] presents results for the EWMD estimator applied to the whole sample of male wages. In this and the subsequent tables the fourth moment matrix has been utilised to correct asymptotic standard errors (reported in parentheses) for the presence of both heteroskedasticity and serial correlation in second moments (see the Appendix). Two goodness of fit measures are reported: the sum of squared

---

[27] The measure ranges from .145 in the case of "perfect mobility" (stochastic independence of earnings in the years delimiting the transition considered), to 1 in the case of "complete immobility" (no changes in ventile ranks).

[28] Each column in the tables containing the estimation results has been numbered and I will refer to each model using such numbers. The same numbers are then used in the corresponding graphs.

residuals and the sum of squared residuals weighted by the inverse of the fourth moment matrix[29, 30].

As a starting point for the analysis, columns 1 to 7 present results obtained without allowing for time varying loading factors within each component. A first thing to note in the table is that the permanent wage component explains about 75% of residual variance in the basic model (model 1, with constant permanent and white noise transitory component), a split similar to that obtained in previous studies (see Dickens [1996]). By letting the transitory component admit some form of serial correlation (models 2, 3 and 4, where the transitory component is AR(1), MA(1) and ARMA(1,1), respectively, while the permanent wage is held constant), we can see how the additional parameter capturing the dispersion of transitory wages in the first year of the panel required by the AR part impacts on the estimate of permanent variance, which tends to be lower in such cases; moreover, in the case of the AR(1) transitory component (model 2) this also affects the estimate of the base transitory variance (σ_ε^2).

When the permanent wage component is allowed to be a linear trend (models 5 to 7), intercepts and slopes of such individual trends negatively covary: the covariance parameter ($\sigma_{\mu\gamma}$) thus captures the compression of differentials which has been observed in figure 2 and indicates the presence of forces equalising wage levels within the permanent wage component.

[29] Moffitt and Gottschalk [1993] adopt the unweighted sum of square residuals, while Dickens [1996] reports the weighted sum of squared residuals: as can be seen from the tables, judgements arising from the two measures are not always in accordance.

[30] Under the null of correct specification, this last measure is distributed as a $\chi 2$ statistic with T(T+1)/2-P (P is the number of estimated parameters) degrees of freedom. Nevertheless, as noted by Dickens [1996], due to large sample size any deviation from the theoretical distribution multiplies up, so that the null is always rejected at conventional levels and the statistic has to be used to compare the fitting performances across models rather than to conduct inference on the specification of the single model.

Moreover, by comparing these models with their counterparts in the left panel of Table 2.4, we could note a sharp increase in the dispersion of individual intercepts (σ_{μ}^2): this indicates that models with a constant permanent component tend to mask the dynamics of permanent wages over the life-cycle, averaging them within the intercept term. This, in turn, affects the size of the initial transitory variance, which is considerably lower (models 2 vs 6).

Another remarkable feature of the results in Table 2.4 is the relatively small dispersion of individual slopes (σ_{γ}^2)[31], which implies that individuals tend to share the same rate of growth of wages through their life-cycles; this evidence seems to be in line with the importance which seniority has traditionally had in the Italian system of wage determination. The estimate of σ_{γ}^2 in model 6 implies that an individual with a slope parameter one standard deviation above the mean will experience a 26% growth of permanent wages over the sample period.

Taking now the transitory component estimates for these models into account, a sharp drop in the AR(1) correlation coefficient can be observed (model 6) with respect to the equivalent specification in the left panel (model 2): in the latter case such parameter has to smooth out the effect of the large initial transitory variance. Finally, it can be seen how no results are reported for the ARMA(1,1) specification of the transitory component: such models didn't converge to a well determined vector of estimated parameters and, in particular, the σ_{ε}^2, σ_0^2 and θ parameters systematically had huge standard errors, leaving the impression that, with a linear permanent

[31] Although the comparison should take into account differences in the definition and measurement of the wage variable, the same coefficient is five to ten times larger in Bourguignon and Morrisson [1983] and Moffitt and Gottschalk [1993] (in this last paper, $\sigma_{\mu\gamma}$ is restricted to be zero).

component, the ARMA(1,1) specification tends to impose an over-parametrization on the data[32].

A way to improve the analysis carried out up to now is to allow for time shifts in the wages covariance structure: as emphasised in Section 1.5, it is important to allow the relative weight of the two wage components to vary over time, in order to capture the effect of different shocks affecting the labour market, which are likely to influence one wage component differently from the other. With this aim, I have estimated models where each wage component includes a year specific loading factor. Such loading factors are specified as a cubic function of time[33] and are normalised to 1 in 1974 for identification. Models are estimated with both versions of the permanent component (constant and linear) and with WN and AR(1) specification of the transitory one.[34] Results for specifications with time-varying loading factors are outlined in the right panel of Table 2.4 (models (8) to (11)) and in Figure 2.5, where the predicted total, permanent and transitory variances of models 9 and 11 are plotted. These predictions are obtained utilising parameter's estimates in the formulas given in (2.4) below, where $E^P(w_{it}w_{is})$ denotes the predicted permanent covariance structure and $E^T(w_{it}w_{is})$ is the predicted transitory covariance structure, while predicted total covariance results from the sum of the two components.

[32] Various attempts in order to avoid such problems were made by minimising the noise in the data dropping n outliers from each tail of the cross-sectional distribution, but without improvements in the results obtained. This strategy was also (this time successfully) pursued in order to solve similar problems in models 10 and 15, where n is equal to 10 and 7.

[33] The choice of the cubic has been dictated by the fact that a quadratic trend is already present in the function to estimate when the permanent component is linear, while linear loading factors could not ensure enough flexibility to fit the empirical covariance structure. Recalling the notation adopted in section 2, $g(t)=1+\alpha_1 t+\alpha_2 t^2+\alpha_3 t^3$ and $h(t)=1+\delta_1 t+\delta_2 t^2+\delta_3 t^3$.

[34] While for the ARMA(1,1) specification the already mentioned over-parametrisation problem was exacerbated the MA(1) specification of the transitory component have not been reported for exposition's compactness: the comparison of results from models (6) and (7) suggests that the restrictive autocorrelation form hypothesised by such models makes them redundant when also AR(1) are estimated.

$$
\begin{aligned}
&E^{P}(w_{it}w_{is}) = g(t)g(s)[\sigma_{\mu}^{2} + \sigma_{\gamma}^{2}(ts) + \sigma_{\mu\gamma}(t+s)] \\
&E^{T}(w_{it}w_{is}) = h(t)h(s)\{\rho^{2}E(v_{it-1}v_{is-1}) + \sigma_{\varepsilon}^{2}[d(1+\theta^{2}) + d_{1}\theta]\} \\
&d = I(|t-s| = 0) \\
&d_{1} = I(|t-s| = 1)
\end{aligned}
\tag{2.4}
$$

By comparing these models with the analogous specifications without time shifters, it can be noticed how the absolute value of $\sigma_{\mu\gamma}$ increases in models 10 and 11 with respect to models 5 and 6: by allowing time specific loading factors, the convergence of wage levels characterising the data from 77 to 82 is not (or is less) averaged with the 74-76 and 87-88 intervals (which present a marked increasing trend), so that its effect on the estimated parameters is stronger. By considering Figure 2.5 it can also be seen that the importance of the transitory component is larger toward the end of the data, a fact which is consistent with the reduction in institutional wage rigidities stressed by the literature on Italy. Such an evidence can also shed light on the behaviour of raw correlation outlined in Figure 2.3: the increasing importance of transitory shocks in the last years of the data means that persistence "washes out" after few years and this is reflected by the fact that the correlation coefficient drops more evidently in the medium than in the short run (Moffitt and Gottschalk [1993]) make a similar point in their analysis of the PSID). Another interesting feature of these graphs is that they allow a better understanding of the effects of the linear specification of the permanent wage component (right panel): this amounts at introducing a higher memory in the permanent variance, which, in this case, doesn't follow the rise in overall variance characterising the last years of the data, and the final increase in transitory variance turns out to be amplified. Overall, while wage convergence operates through the permanent wage, the re-opening of differentials is also determined by the increased volatility. However, in the light of the descriptive mobility analysis of Figure 2.4, permanent wages convergence doesn't seem to have induced any increase in the degree of hierarchical mobility.

The fact that a negative value of $\sigma_{\mu\gamma}$ doesn't show up in variations of quantile rankings casts doubts on a human capital interpretation of the aggregate wage convergence process. On the other hand, the egalitarian wage indexation system effective in the first part of the period considered could be consistent with a generalised convergence of the wage distribution which leaves the relative wage hierarchy unaltered. In an attempt to shed some light on this institutional interpretation, the analysis has been extended by estimating separate models for blue and white collar workers. The assumption behind this exercise is that if market forces also contributed to the convergence process, the permanent mobility effect should still be present after conditioning on occupation, institutional factors mainly operating between occupations. Moreover, if convergence is determined by the remuneration of human capital investments along the lines of the Mincerian on-the-job training model, it should still be effective within occupations (see, for example, Lillard and Weiss [1979]).

Results from this experiment are reported in Table 2.5 (models (12) to (19)) and in Figures 2.6 and 2.7. The striking fact emerging from models 18 and 19 is that the $\sigma_{\mu\gamma}$ parameter is positive for white collar workers, thus implying a life-time divergence of wage levels which contrasts with the permanent wage convergence singled out in Table 4 for the whole sample. A positive $\sigma_{\mu\gamma}$ could be consistent with a signalling/matching framework, in which more productive workers earn higher initial wages according to their observable abilities and experience faster growth due to the increasing value of the match; also this evidence is consistent with the diffusion of individual wage premia in the 1980s, as reported by Dell'Aringa and Lucifora [1994], and mirrors the increases in autocorrelation and immobility in this sub-sample shown by the descriptive analysis of Figures 2.3 and 2.4.

The $\sigma_{\mu\gamma}$ parameter is instead negative for blue collar data, where the empirical wage dispersion plotted in Figure 2.2 decreased until 1987. Another fact to note in the Table is the higher dispersion of residual wage profiles' slopes (σ_{γ}^{2}) in white collars data. Finally a higher base variance of the transitory component (σ_{ε}^{2}) can be noted in blue collars data. The inspection of Figure 6 reveals that this last finding is mainly due to the dynamics in the extreme years of the data: if we recall that the empirical trend to fit is decreasing from 77 to 87 and sharply increasing from 74 to 76 and from 87 to 88 (Fig. 2.2), it is clear how the permanent component could not pick up such drastic changes, so that they have to be recorded as transitory. Moreover, if we specify a more persistent permanent variance (Fig. 2.6 right panel), the transitory component accounts for all the wage dispersion of 1988. In general, what can be observed in figure 6 is a lagged and persistent reaction of the permanent variance profile to the evolution of the total variance one. The situation is different in white collar's data (Fig. 2.7), where the permanent variance profile closely follows that of total variance and tends to increase its predominance toward the end of the data. It is worth noting how here the linear specification of the permanent component doesn't change the permanent variance profile in the last years of the data as it did for blue collars' data: the empirical variance for white collars increases since 1984, so that there's enough time to incorporate the increasing trend within the permanent wage component.

Results obtained by conditioning on occupation could be driven by the fact that such characteristic is not time-invariant: as time elapses, changes of occupational status may take place and this could lead to some misleading conclusions. In particular, such changes could affect workers in the upper tail of the origin wage distribution, who, as a further development of their career, reach a higher occupational status. In the light of these considerations, results such as the convergence of permanent wages observed for the blue collars distribution should be considered with caution. In order to check the robustness of results to the effects of between occupations mobility, models have been re-estimated imposing the constancy of the occupational status. Thus, rather than considering the wage distribution of workers with a given occupation in a given year, the focus has been pointed to the occupation a worker has the first year he appears in the panel, which has been assumed as his time-invariant occupational status. Hence, the models estimated also take into account the effects of the occupational career on the individual wage profile. Results are contained in the right panel of Table 2.5 (models (20) and (21)) and show that while both the intercepts and slopes variances tend to increase (not surprisingly given the design of the occupational classification), the sign of the convergence parameters remains unchanged, thus confirming the evidence arising from the previous analysis.

From a permanent mobility viewpoint, results from Table 2.5 suggest that the convergence of permanent wage levels observed in the overall wage distribution was to a great extent due to the effects of the wage indexation system, being absent in white collars data, where the $\sigma_{\mu\gamma}$ parameter is positive. Of course, this evidence is not enough to conclude that human capital investments have not been remunerated for white collar workers: it may well be that other factors, such as an increase in the demand for skilled labour, were operating in this period and that their effectiveness was enhanced by the relaxation of the compressing forces outlined above, thus offsetting any converging trend.

On the other hand, we observe a convergence of the blue collar wage distribution which substantially covers the whole period. Such convergence operates through the permanent wage component ($\sigma_{\mu\gamma}<0$), so that one would be tempted to conclude that the blue collars wage distribution behaved according to the human capital paradigm. Nevertheless, if we carefully observe the dynamics depicted in the right panel of Figure 2.6, we can note that the permanent wage variance reacts with a certain lag to the development of total variance: in particular, permanent is increasing in 1976-77, while total variance is decreasing, and the opposite is true in the 85-87 interval. Hence, the evidence of permanent mobility could be due to the lagged reaction of the permanent component, which makes the effect of the egalitarian wage indexation system more persistent.

2.5. Summary and conclusions

Results arising from the above analysis clearly underline some features of the dynamics of the Italian wage distribution over the 1974-88 period.

First of all, a clear reversion in inequality trends has been detected, with a declining phase characterising the data from the end of the 1970s to the first half of the 1980s, and a reopening of wage differentials (especially between blue and white collar workers and within white collars) thereafter. Such evidence is not new, being reported by the existing literature on the Italian wage distribution, and many authors agree in attributing it, at least to a large extent, to the egalitarian system of wage indexation which was effective in the first part of the period considered. By taking advantage of the longitudinal structure of the data utilised, the paper has also looked at indicators of autocorrelation and mobility. While for the overall wage distribution and for blue collar workers there's a tendency for autocorrelation to decrease (especially in the medium term) and for immobility to be stable after the u-turn in inequality, white collar data display increments in both measures of persistence over the last years of the sample.

The econometric analysis has focused on the estimation of the parameters of the joint wage distribution and, in particular, on the investigation of the permanent and transitory components of inequality and mobility. Results show how a convergence of permanent wage levels is indicated by the data, and it has been interpreted as reflecting the equalising institutional factors at work. Moreover, the relative weight of the transitory wage component is found to increase toward the end of the sample period, signalling an increased vulnerability of wages to transitory shocks during and thus a higher uncertainty of labour incomes. Finally, by extending the analysis within occupational groups, permanent wage divergence has been found in white collar data, suggesting that institutional forces, mainly effective between occupations, were in fact an important factor in determining the observed overall convergence.

From a policy viewpoint, these results imply a certain inability of the Italian labour market in stimulating human capital accumulation and, given that the effectiveness of equalising institutional factors has been reduced in recent years, suggest that great attention should be devoted to the wage careers of workers at the bottom end of the wage distribution.

2. Tables and Graphs

Table 2.1: The structure of the INPS panel

	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
1974	7291	7055	6985	6680	6859	6863	6814	6497	6194	5798	5768	5735	4842	4855	3754
1975		9083	8762	8384	8524	8523	8502	8187	7844	7395	7431	7345	5895	5750	4537
1976			9619	8965	9021	8996	8998	8669	8308	7834	7856	7773	6259	6127	4780
1977				9705	9129	9098	9084	8759	8375	7960	7916	7813	6314	6133	4837
1978					10630	10232	10086	9706	9302	8766	8720	8614	7033	6739	5241
1979						11089	10548	10154	9753	9165	9111	8976	7372	7043	5475
1980							11621	10767	10306	9637	9613	9464	7761	7422	5772
1981								11737	10875	10142	10082	9631	7907	7610	5948
1982									11721	10334	10630	9958	8190	7835	6098
1983										11438	10122	9466	7679	7304	5978
1984											11914	10678	8717	8185	6441
1985												12275	8941	8404	6669
1986													10277	8184	6429
1987														10623	8334
1988															10059

Note: each cell gives the number of individuals whose wage residuals are used in the computation of the corresponding cell of the wage covariance matrix in table 3

Table 2.2: Sample proportions by years and selected workers' characteristics (upper panel) and descriptive statistics of the cross-sectional log-wage distribution (lower panel)

	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
%															
Blue collars	69.1	71.74	79.22	85.84	70.46	70.14	70.15	71.61	72.24	71.69	73.1	71.27	70.73	70.97	70.06
White collars	30.13	27.57	20.18	21.68	28.66	28.92	28.72	27.42	26.82	27.28	25.72	27.23	27.85	27.29	27.86
cohort 1923-24	1.23	1.18	1.15	1.09	1.12	1.10	1.10	1.15	1.16	1.15	1.18	1.13	1.06	0.96	0.88
cohort 1925-29	8.64	8.24	8.17	8.04	7.96	7.76	7.63	7.62	7.71	7.34	7.55	7.51	7.03	7.40	7.15
cohort 1930-34	19.27	19.20	18.60	18.48	18.25	17.95	18.01	17.77	17.93	17.78	18.05	17.60	17.66	18.08	17.34
cohort 1935-39	22.18	22.02	21.59	21.49	21.46	21.41	21.14	21.07	20.83	20.41	20.85	20.74	20.19	20.78	20.65
cohort 1940-44	21.52	21.17	21.29	20.96	21.10	20.89	21.29	21.13	20.89	21.34	20.68	20.88	20.86	20.71	20.95
cohort 1945-49	20.85	21.39	21.86	22.33	22.19	22.53	22.49	22.71	22.63	23.09	22.81	23.02	23.84	22.96	23.44
cohort 1950-51	6.31	6.80	7.33	7.61	7.92	8.36	8.34	8.55	8.86	8.88	8.89	9.12	9.36	9.11	9.59
Mean	5.70	5.60	5.67	5.67	5.71	5.74	5.78	5.80	5.79	5.82	5.81	5.85	5.85	5.89	5.89
St.Dev.	0.559	0.577	0.612	0.580	0.549	0.529	0.504	0.464	0.439	0.465	0.456	0.490	0.461	0.470	0.560

Table 2.3: The covariance matrix of wages - Whole sample

	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988
1974	.310	.266	.259	.251	.224	.205	.199	.175	.161	.156	.161	.191	.173	.197	.225
	(.011)	(.010)	(.008)	(.011)	(.007)	(.007)	(.007)	(.006)	(.006)	(.006)	(.006)	(.007)	(.007)	(.008)	(.010)
1975	.832	.330	.286	.278	.260	.240	.230	.202	.186	.189	.184	.216	.192	.205	.239
		(.010)	(.007)	(.009)	(.007)	(.007)	(.006)	(.006)	(.005)	(.006)	(.006)	(.007)	(.006)	(.007)	(.010)
1976	.759	.813	.374	.283	.268	.248	.238	.211	.189	.198	.184	.218	.185	.201	.229
			(.008)	(.007)	(.007)	(.006)	(.006)	(.005)	(.005)	(.005)	(.005)	(.006)	(.005)	(.007)	(.008)
1977	.778	.837	.800	.334	.271	.250	.237	.208	.193	.199	.187	.218	.192	.193	.221
				(.012)	(.008)	(.007)	(.007)	(.006)	(.006)	(.006)	(.006)	(.007)	(.006)	(.007)	(.009)
1978	.733	.826	.801	.857	.300	.254	.237	.212	.194	.200	.191	.223	.196	.187	.213
					(.007)	(.006)	(.006)	(.005)	(.005)	(.005)	(.005)	(.006)	(.006)	(.006)	(.008)
1979	.697	.791	.769	.819	.878	.279	.227	.201	.183	.192	.182	.212	.181	.175	.202
						(.007)	(.006)	(.005)	(.005)	(.005)	(.005)	(.006)	(.005)	(.006)	(.007)
1980	.708	.797	.773	.814	.860	.854	.253	.199	.181	.190	.178	.208	.174	.176	.199
							(.006)	(.005)	(.004)	(.005)	(.005)	(.005)	(.005)	(.006)	(.007)
1981	.677	.760	.747	.779	.834	.821	.853	.215	.176	.182	.172	.188	.166	.151	.171
								(.005)	(.004)	(.004)	(.004)	(.005)	(.005)	(.005)	(.006)
1982	.660	.739	.704	.763	.809	.790	.822	.867	.192	.173	.165	.174	.157	.143	.161
									(.005)	(.004)	(.004)	(.004)	(.004)	(.004)	(.005)
1983	.603	.709	.697	.740	.788	.782	.813	.847	.851	.216	.179	.189	.157	.141	.155
										(.006)	(.004)	(.005)	(.004)	(.004)	(.005)
1984	.633	.702	.662	.711	.764	.757	.777	.816	.824	.846	.208	.182	.165	.156	.168
											(.005)	(.004)	(.004)	(.005)	(.006)
1985	.702	.770	.731	.771	.833	.821	.843	.830	.811	.830	.817	.239	.183	.179	.198
												(.005)	(.005)	(.005)	(.006)
1986	.677	.726	.657	.724	.777	.747	.753	.778	.777	.735	.787	.812	.212	.169	.189
													(.006)	(.005)	(.006)
1987	.755	.759	.702	.711	.727	.709	.745	.697	.694	.648	.730	.782	.782	.220	.218
														(.006)	(.006)
1988	.724	.743	.671	.684	.695	.684	.710	.660	.659	.598	.662	.724	.738	.834	.312
															(.008)

Notes: asymptotic standard errors in parentheses; correlation coefficients below the diagonal

Table 2.4: Estimates for the whole sample

	μ_i+WN (1)	μ_i+AR(1) (2)	μ_i+MA(1) (3)	μ_i+ ARMA(1,1) (4)	$\mu_i+\gamma_i t$+ WN (5)	$\mu_i+\gamma_i t$+ AR(1) (6)	$\mu_i+\gamma_i t$+ MA(1) (7)	$g(t)\mu_i$+ h(t)WN (8)	$g(t)\mu_i$+ h(t)AR(1) (9)	$g(t)(\mu_i+\gamma_i t)$+ h(t)WN (10)	$g(t)(\mu_i+\gamma_i t)$+ h(t)AR (1) (11)
$\sigma^2\mu$	0.199	0.150	0.196	0.137	0.275	0.266	0.268	0.224	0.206	0.201	0.202
	(0.0046)	(0.0109)	(0.0047)	(0.0142)	(0.0076)	(0.0077)	(0.0076)	(0.0100)	(0.0101)	(0.0085)	(0.0103)
$\sigma^2\gamma$					0.0005	0.0003	0.0004			0.0006	0.0003
					(0.00004)	(0.00004)	(0.00004)			(0.00003)	(0.0001)
$\sigma\mu\gamma$					-0.0070	-0.0062	-0.0063			-0.0104	-0.0094
					(0.0004)	(0.0004)	(0.0004)			(0.0004)	(0.0006)
$\sigma^2\varepsilon$	0.067	0.013	0.061	0.065	0.057	0.059	0.057	0.082	0.076	0.085	0.155
	(0.0016)	(0.0010)	(0.0015)	(0.0020)	(0.0015)	(0.0015)	(0.0014)	(0.0056)	(0.0100)	(0.0041)	(0.0125)
$\sigma^2 0$		0.185		0.200		0.045			0.103		0.097
		(0.0094)		(0.0121)		(0.0083)			(0.0068)		(0.0069)
ρ		0.870		0.912		0.297			0.472		0.801
		(0.0158)		(0.0127)		(0.0259)			(0.0170)		(0.0140)
θ			0.366	-0.714			0.262				
			(0.0190)	(0.0211)			(0.0176)				
$\alpha 1$								0.068	0.082	0.209	0.253
								(0.0088)	(0.0098)	(0.0168)	(0.0164)
$\alpha 2$								-0.017	-0.018	-0.037	-0.048
								(0.0014)	(0.0015)	(0.0030)	(0.0029)
$\alpha 3$								0.0009	0.001	0.002	0.003
								(0.0001)	(0.0001)	(0.0002)	(0.0002)
$\delta 1$								0.161	0.121	-0.068	-0.442
								(0.0272)	(0.0475)	(0.0135)	(0.0136)
$\delta 2$								-0.062	-0.055	0.001	0.049
								(0.0054)	(0.0092)	(0.0026)	(0.0025)
$\delta 3$								0.0029	0.003	0.000	-0.002
								(0.0003)	(0.0004)	(0.0001)	(0.0001)
ssr	0.1526	0.098	0.1465	0.067	0.0791	0.075	0.0767	0.0399	0.029	0.02572	0.028
χ^2	3644.14	3071.78	5740.11	2125.29	2623.19	2891.99	3319.71	2698.81	2561.59	2172.31	3114.31

Notes: Asymptotic standard errors in parentheses. Columns are labelled according to the model specified for w_{it} (as defined in the text). Estimates shown refer to (in descending order): covariance **structure of the core permanent wage** (from $\sigma^2\mu$ to $\sigma\mu\gamma$), covariance structure of the core transitory wage (from $\sigma^2\varepsilon$ to θ), coefficients of g(t) (the αs) and h(t) (the δs).

Table 2.5: Estimates by workers' occupation

	Blue Collars				White Collars				Blue Collars	White Collar
	$g(t)\mu_i$+ h(t)WN (12)	$g(t)\mu_i$+ h(t)AR(1) (13)	$g(t)(\mu_i+\gamma_i t)$+ h(t)WN (14)	$g(t)(\mu_i+\gamma_i t)$+ h(t)AR (1) (15)	$g(t)\mu_i$+ h(t)WN (16)	$g(t)\mu_i$+ h(t)AR(1) (17)	$g(t)(\mu_i+\gamma_i t)$+ h(t)WN (18)	$g(t)(\mu_i+\gamma_i t)$+ h(t)AR (1) (19)	$g(t)(\mu_i+\gamma_i t)$+ h(t)AR (1) (20)	$g(t)(\mu_i+\gamma_i t)$+ h(t)AR (1) (21)
$\sigma^2\mu$	0.140	0.117	0.127	0.095	0.180	0.172	0.221	0.111	0.131	0.219
	(0.0094)	(0.0081)	(0.0075)	(0.0061)	(0.0116)	(0.0128)	(0.0130)	(0.0475)	(0.0096)	(0.0137)
$\sigma^2\gamma$			0.0005	0.0003			0.0091	0.0056	0.0004	0.0082
			(0.00003)	(0.00002)			(0.0038)	(0.0035)	(0.00003)	(0.0024)
$\sigma\mu\gamma$			-0.0078	-0.0057			0.0279	0.0378	-0.0075	0.0299
			(0.0004)	(0.0004)			(0.0073)	(0.0108)	(0.0005)	(0.0057)
$\sigma^2\varepsilon$	0.117	0.112	0.122	0.104	0.071	0.045	0.027	0.024	0.068	0.089
	(0.0072)	(0.0121)	(0.0073)	(0.0108)	(0.0077)	(0.0156)	(0.0073)	(0.0090)	(0.0098)	(0.0172)
$\sigma^2 0$		0.146		0.135		0.071		0.128	0.134	0.021
		(0.0091)		(0.0072)		(0.0104)		(0.0438)	(0.0088)	(0.0136)
ρ		0.451		0.549		0.594		0.626	0.413	0.513
		(0.0154)		(0.0203)		(0.0335)		(0.0312)	(0.0195)	(0.0361)
$\alpha 1$	0.099	0.112	0.297	0.306	0.035	0.026	-0.144	-0.144	0.243	-0.145
	(0.0168)	(0.0193)	(0.0274)	(0.0311)	(0.0145)	(0.0159)	(0.0232)	(0.0269)	(0.0259)	(0.0155)
$\alpha 2$	-0.018	-0.017	-0.046	-0.046	-0.012	-0.011	0.009	0.009	-0.034	0.009
	(0.0022)	(0.0026)	(0.0047)	(0.0056)	(0.0025)	(0.0028)	(0.0031)	(0.0035)	(0.0042)	(0.0021)
$\alpha 3$	0.0007	0.0005	0.0030	0.0034	0.0007	0.0007	-0.0002	-0.0002	0.0022	-0.0001
	(0.0001)	(0.0001)	(0.0003)	(0.0004)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0003)	(0.0001)
$\delta 1$	0.164	0.137	0.106	0.025	-0.057	0.068	0.273	0.238	0.037	-0.104
	(0.0248)	(0.0390)	(0.0238)	(0.0306)	(0.0337)	(0.1031)	(0.1027)	(0.1283)	(0.0430)	(0.0442)
$\delta 2$	-0.061	-0.054	-0.048	-0.025	0.016	-0.016	-0.047	-0.041	-0.020	0.007
	(0.0049)	(0.0073)	(0.0047)	(0.0055)	(0.0062)	(0.0154)	(0.0164)	(0.0189)	(0.0073)	(0.0080)
$\delta 3$	0.0029	0.0025	0.0023	0.0017	-0.0016	0.0008	0.0021	0.0018	0.0013	-0.0001
	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0007)	(0.0007)	(0.0008)	(0.0003)	(0.0004)
ssr	0.0563	0.041	0.0486	0.026	0.0371	0.018	0.0269	0.018	0.022	0.025
χ^2	2586	3552.41	2191.96	3755.92	1143.47	855.02	1124.74	894.437	2553.92	1043.58

Notes: Asymptotic standard errors in parentheses. Columns are labelled according to the model specified for w_{it} (as defined in the text). Estimates shown refer to (in descending order): covariance structure of the core permanent wage (from $\sigma^2\mu$ to $\sigma\mu\gamma$), covariance structure of the core transitory wage (from $\sigma^2\varepsilon$ to θ), coefficients of g(t) (the αs) and h(t) (the δs). In columns 20 and 21 the occupational classification held in the first year in which the worker is observed is maintained constant in all the panel waves

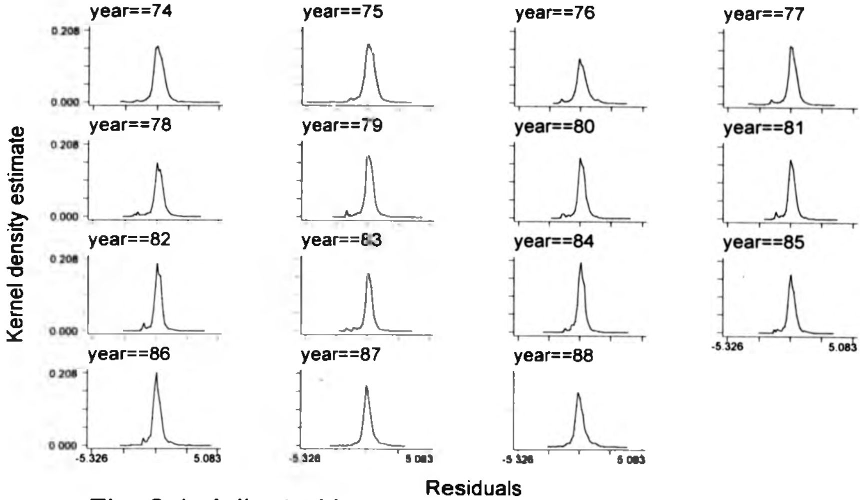

Fig. 2.1: Adjusted log-wage cross-sectional distributions

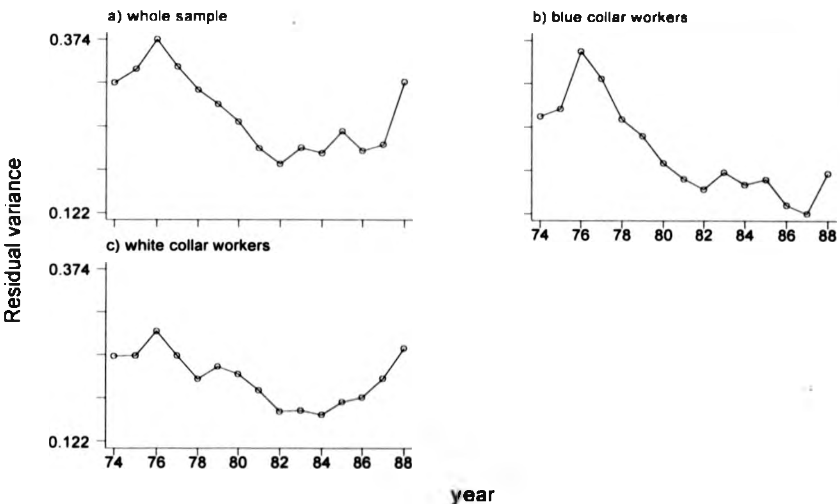

Fig. 2.2: Adjusted log-wage variances

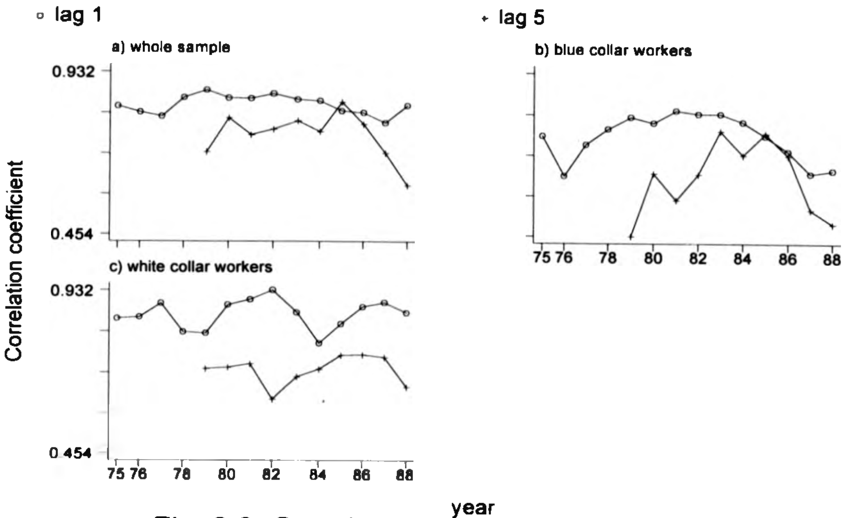

Fig. 2.3: Correlation coefficients at lag 1 and 5

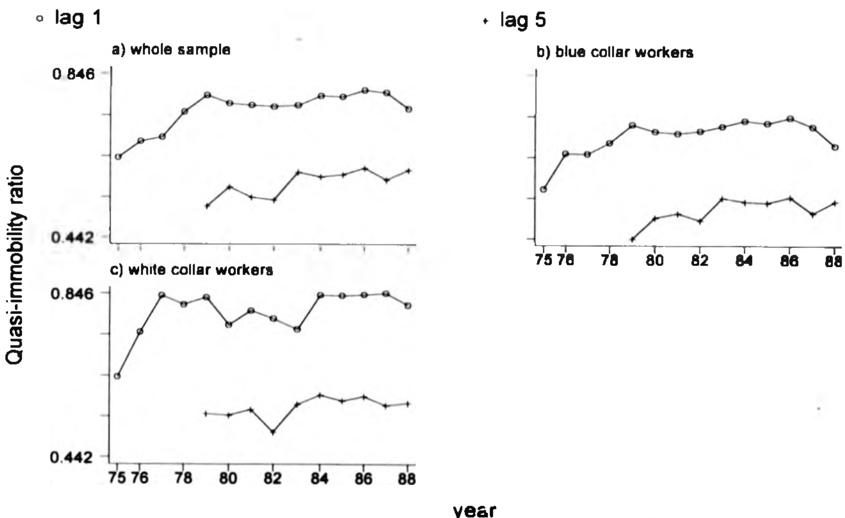

Fig. 2.4: Ventile quasi-immobility ratios at lag 1 and 5

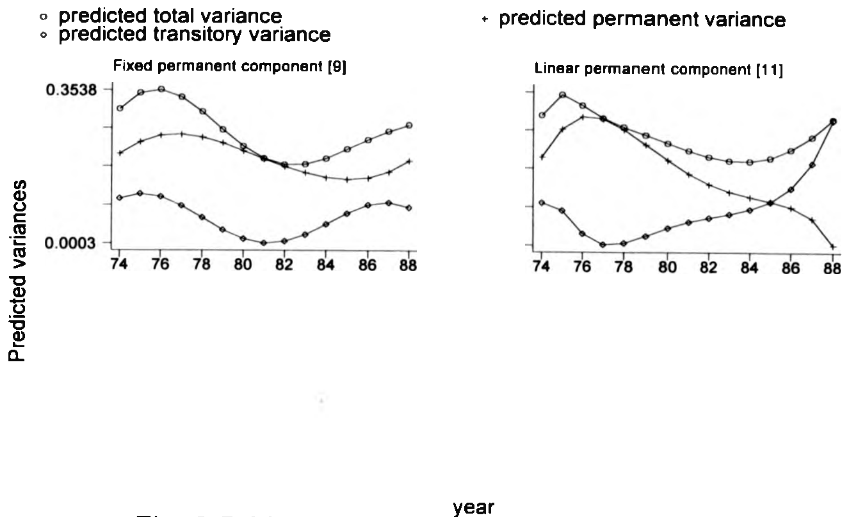

Fig. 2.5: Variance decomposition - Whole sample

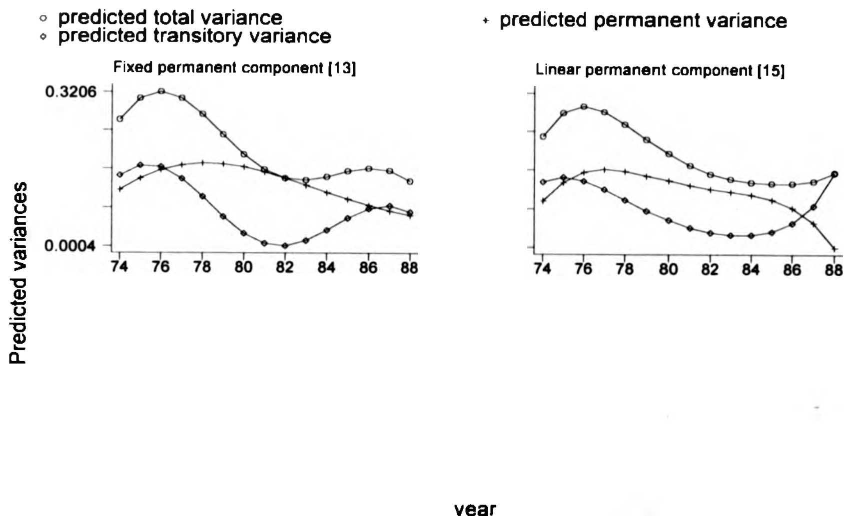

Fig. 2.6: Variance decomposition - Blue collar workers

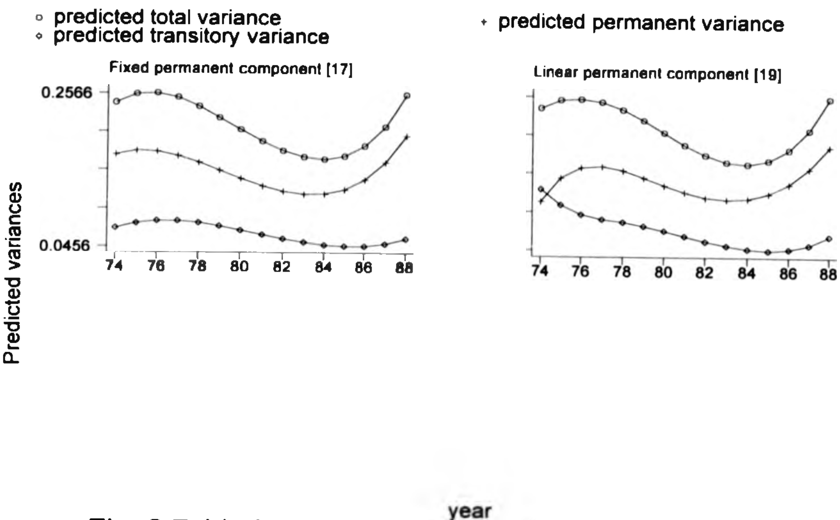

Fig. 2.7: Variance decomposition - White collar workers

Appendix A2: Minimum distance estimation of variance components models for the wage covariance structure

This Appendix presents the method utilised in the estimation of the (adjusted) wage covariance structure and of the variance components models of Chapters 2 and 3 and the STATA codes written for its implementation. The procedure is the one set out in Chamberlain [1984] and Abowd and Card [1989] and extended to unbalanced panels by Dickens [1996]; see also Crepòn and Mairesse [1997] for a general overview of the method.

Our data set consists of information on wages and explanatory variables on N individuals observed over a time span of length T, with individual records which may be missing in a given year. Let y_{it} denote the real wage of worker i in year t.

As a first step in the procedure, raw wages are regressed on a set of life-cycle, business-cycle and productivity growth indicators in order to abstract from such effects in the variance decomposition analysis; in particular, this is done with an OLS regression pooling the panel waves:

$$\begin{aligned} \log y_{it} &= x'_{it}\beta + u_{it} \\ u_{it} &\sim WN(0,\sigma_u^2) \end{aligned} \qquad \text{(A2.1)}$$

Let w_{it} denote the residuals from this first stage regression. The following step in the procedure requires estimation of the second moments of such residuals, which will be given by the sample average of individual second moments. To form the second moments matrix (W_i) for individual i, define w_i as the vector of residuals for individual i: $w_i' = (w_{i1},\ldots,w_{iT})$, with w_{is}=0 if worker i is not observed in year s. Then: $W_i = w_i w_i'$. Next define a dummy variable $d_{it} = \mathrm{I}(w_{it} \neq 0)$, and, correspondingly, the vectors $d_i' = (d_{i1}, \ldots, d_{iT})$ and the matrices $D_i = d_i d_i'$. Then, the empirical second moments matrix (C) will be given by the element by element ratio of the two matrices W and D, with $W = \sum_i W_i$ and analogously for D.

Let m=vech(C), i.e. the T(T+1)/2×1 vector containing the distinct elements of C, and m_i=vech(W_i). The main point for asymptotic properties of the estimator is that independence of the w_i implies independence of m_i. Chamberlain [1984] shows that (under fairly general conditions) $m_i \sim N(m, V)$, where V is the fourth moments matrix.

An estimate of V is given by the empirical fourth moments matrix. Define p=vech(D) and $\tilde{m}_i = m_i - m$. Then V may be estimated as the element by element ratio of the two matrices $\tilde{M}$ and P, where $\tilde{M} = \sum_i \tilde{m}_i \tilde{m}_i'$ and $P = pp'$.

The STATA code `covax.do` which implements estimation of the second and fourth moments matrices is reported below. The program uses as input the file `matres.dta` which contains the T columns of residuals from the first stage regression in (A2.1) (the variables named `res*`, where * goes from 74 to 88, i.e. the code refers to the data set of Chapter 2) and the INPS identification number (the variable `ind`) for N individual records.

```
use matres                                        /* covax.do */

capture program drop codda  /*fills in the gaps of the
                            unbalanced panel*/

program define codda
      version 5.0
      local j=74
      while `j'<=88 {
            recode res`j' .=0
            local j=`j'+1
      }
end
codda


capture program drop cprod  /*creates cross-products of
                                  wit*/
program define cprod
      version 5.0
      local i=74
      while `i'<=88 {
            local j=`i'
            while `j'<=88 {
                  gen res`i'`j'=res`i'*res`j' /*element
                                               of Wi*/
                  local j=`j'+1
            }
            gen d`i'=res`i'~=0     /*generates the dummy for
                                   presence in sample*/
            drop res`i'
            local i=`i'+1
      }
end
cprod

matrix accum bop=d*          /*provides the description of
                                   panel structure*/
```

```
capture program drop cdum        /*cross-products of
                                       dummies*/
program define cdum
      version 5.0
      local i=74
      while `i'<=88 {
            local j=`i'
            while `j'<=88 {
                  gen d`i'`j'=d`i'*d`j' /*element of Di*/
                  local j=`j'+1
            }
            drop d`i'
            local i=`i'+1
      }
end
cdum

capture program drop covax /*second and (half of) fourth
                                 moments*/
program define covax
      version 5.0
      local i=74
      while `i'<=88 {
            local j=`i'
            while `j'<=88 {
                  egen s`i'`j'=sum(res`i'`j')
                                         /*element of W*/
                  egen t`i'`j'=sum(d`i'`j')
                                         /*element of D*/
                  gen m`i'`j'=s`i'`j'/t`i'`j'
                                         /*element of M*/
                  quietly gen dev`i'`j'=(m`i'`j'-
                  res`i'`j')/(t`i'`j') if res`i'`j'~=0
                  /*standardised deviations of individual c-
            p from mean c-p*/
                  quietly recode dev`i'`j' .=0
                  drop res`i'`j' s`i'`j' d`i'`j'  t`i'`j'
                  local j=`j'+1
            }
            local i=`i'+1
      }
      sort ind
end
covax
```

```
save tempo,replace

use tempo
matrix accum fm=dev*  ,noc /*gets the fourth moments
                               matrix*/
drop _all
svmat fm,n(col)
save fmp,replace      /*and saves it*/
use tempo
mkmat m* in 1/1, matrix (m) /*gets the vector of second
                               moments*/
matrix mm=m'
drop _all
svmat mm,n(m)
save mmp,replace      /*and saves it*/
```

The variance decomposition analysis of Chapters 2 and 3 is performed by fitting the theoretical covariance structure implied by the hypothesised model of wage residuals to the empirical covariance matrix contained in m. The theoretical covariance structure will be, in general, a non linear function of the parameters of the model, say f(b). As Chamberlain [1984] shows, the restrictions implied by the theoretical model can be imposed by

$$\min_b \sum_i [m_i - f(b)]' A [m_i - f(b)] \qquad \text{(A2.2a)}$$

or, equivalently, by

$$\min_b [m - f(b)]' A [m - f(b)] \qquad \text{(A2.2a)}$$

where A is some suitable weighting matrix.

The minimum distance estimator $\hat{b}$ is shown to be consistent for b with asymptotic covariance matrix given by $V(\hat{b}) = (G'AG)^{-1} G'AVAG(G'AG)^{-1}$, where G= $\frac{\delta f(b)}{\delta b}\Big|_{b=\hat{b}}$ is the gradient matrix evaluated at the solution of the minimisation problem (see Chamberlain [1984]).

The choice of A generates a class of minimum distance estimators. In particular, Chamberlain [1984] suggests the adoption of $A=V^{-1}$, which yields the optimum minimum distance (OMD) estimator. Recently, Altonji and Segal [1996] have provided Monte Carlo evidence showing that correlation between second and fourth moments could bias the OMD in finite samples; they argue in favour of an equally weighted minimum distance estimator (EWMD), where $A=I$.

A test for the estimated model against the alternative of an unrestricted covariance structure is given by the sum of squared residuals weighted by the inverse of the fourth moments matrix, which, under the null of correct model specification, has a $\chi 2$ distribution with T(T+1)/2-q degrees of freedom, where q is the dimension of *b*.

As pointed out in Section 2.2, the derivation of f(b) poses a problem due to the initial conditions of autoregressive stochastic processes: in particular, the process cannot be deemed to have started in the infinite past (as customary in time series analysis) and the variance of the initial conditions has to be explicitly modelled within the covariance structure parameters. For illustrative purposes, below I show the implications of this initial conditions problem in the case of an AR(1) process:

$$\begin{aligned} v_{it} &= \rho v_{it-1} + \varepsilon_{it} \\ \varepsilon_{it} &\sim WN(0,\sigma_\varepsilon^2) \end{aligned} \tag{A2.3}$$

which, tracing the recursion back to the initial year of data, can be written as:

$$v_{it} = \rho^t v_{i0} + \sum_{j=0}^{t-1} \rho^j \varepsilon_{it-j} \tag{A2.4}$$

After some algebra we get the autocovariance function which takes into account the variance of the initial conditions of the process (σ_0^2):

$$E(v_{it}v_{is}) = \begin{cases} \rho^{s}\sigma_0^2 & \text{if} \quad s \geq t = 0 \\ \rho^{(t+s)}\sigma_0^2 + \rho^{|t-s|}\dfrac{1-\rho^{2t}}{1-\rho^{2}}\sigma_\varepsilon^2 & \text{if} \quad s \geq t \geq 1 \end{cases}$$

$$s,t = 0,\ldots,T-1 \qquad \text{(A2.5)}$$
$$s \geq t$$

The EWMD estimator of the covariance structure has been implemented using STATA's non linear least squares routine (nl) and extending it to obtain the asymptotic robust standard errors; below I illustrate the code for a simple process with constant permanent component and AR(1) transitory component:

$$\begin{aligned} w_{it} &= \mu_i + v_{it} \\ v_{it} &= \rho v_{it-1} + \varepsilon_{it} \\ \mu_i &\sim (0,\sigma_\mu^2) \\ v_{i0} &\sim (0,\sigma_0^2) \\ \varepsilon_{it} &\sim WN(0,\sigma_\varepsilon^2) \end{aligned} \qquad \text{(A2.6)}$$

```
/*konar.do*/
set matsize 800
use fmp                        /*loads fourth moments colums*/
mkmat dev*,matrix(fm)          /*forms the matrix*/
drop _all
use mmp                   /*loads the vector of second
                          moments*/

capture drop resi

capture program drop nlkonar
program define nlkonar
     version 5.0
     if "`1'"=="?" {
          global S_1 "b0 b3 b30 b4"   /*declares and
                                     initialises
                                     parameters*/
          global b0=.2                /*σ²_μ*/
          global b3=.05               /*σ²_ε*/
          global b30=.02              /*σ²_0*/
          global b4=.8                /*ρ*/
          exit
     }
     replace `1'=$b0 +($b30*$b4^s)*dr0+($b30*$b4^(t+s) +
     ((1-$b4^(2*t))/(1-$b4^2))* $b3*$b4^tds)*(1-dr0)
     /*autocovariance function*/

end

quietly nl konar mmp t s tds, leave
/*the leave option gets the columns of the gradient matrix*/
/*t and s are time indices for the rows and columns of the
matrix respectively, tds=|t-s|,dr0=I(t=0)*/
nlpred resi, resid                  /*residuals needed to
                                    compute the chi2 stat*/
mkmat b0 b1 b2 b3 b4, matrix (g) /*forms the gradient
                                        matrix*/
matrix gg=g'*g
matrix gginv=inv(gg)
matrix pp1=g'*fm
matrix pp2=pp1*g
matrix pp3=gginv*pp2

matrix vp=pp3*gginv          /*V(b̂) = (G'G)^-1 G'VG(G'G)^-1*/
mkmat resi, matrix(resi)
matrix fminv=inv(fm)
matrix chi1=resi'* fminv

matrix chi=chi1*resi         /*χ2 = (m - f(b̂))'V^-1(m - f(b̂))*/
```

Chapter 3

Male wage inequality dynamics: permanent changes or transitory fluctuations?

3.1 Introduction

This Chapter presents an extension of the analytical framework of Chapter 2 in several directions. A first relevant feature of the empirical models utilised resides in the modelling of the time-varying loading factors on each wage component. In particular, following the work of Dickens [1996], I relax the polynomial specification of the loadings to allow them to freely vary in each sample year, i.e. each wage component will be characterised by T-1 (where T is the time dimension of the data) time shifters to be estimated. Such a specification presents the advantage of not imposing any a priori functional form on the loading factors, so that the extent to which changes in aggregate differentials are permanent or transitory can be better assessed.

A second innovation with respect to the analysis of Chapter 2 is that (similarly to Baker [1997]) the random growth specification of the permanent wage component will be compared with a random walk model, the assumption underlying the two specifications being rather different. As we will see below, conversely to Baker's results, the INPS data are, in various circumstances, favourable to the random walk assumption, thus pointing to a context in which individual permanent wage dynamics present a high degree of persistence and, rather than following person specific profiles as would be predicted by (e.g.) human capital theories of wage dynamics, move erratically around their long-run level.

Thirdly, greater attention will be paid to the role of observable workers characteristics. This will be done in two steps. First, the variance decomposition analysis will be extended to wages adjusted also for the effect of observable characteristics, i.e. within the cells defined by such explanatory variables. Then, after performing within group analyses on sub-samples defined by such characteristics (as in Chapter 2), a method for conditioning the covariance structure on such observables will be introduced, thus providing a unified framework for assessing the impact of the variables of interest on the wage covariance structure and providing additional insights into its dynamics.

Finally, a noticeable difference is given by the data set. As mentioned in the Introduction to the Thesis, only a "small" INPS draw was available at the time this research began, and this is the data set utilised in the previous Chapter. Afterwards, a second (unbalanced) draw has been made available, which is larger in size (with the estimation sample which is more than triple the one used in Chapter 2) and covers a more recent period of time, in particular including the years in which automatisms in wage indexation have been completely removed.

The chapter is structured as follows. Section 3.2 discusses the analytical framework, while Section 3.3 describes the "large" INPS sample. Section 3.4 provides some descriptive statistics on the covariance structure, while the EWMD estimation results are presented in Section 3.5. Section 3.6 draws some conclusions.

3.2 The analytical framework

The model specification in this chapter has two main differences compared to the restrictions imposed on the covariance structure of Chapter 2. Let us define our model of adjusted log-wages as :

$$w_{it} = g(t)w_{it}^{P} + h(t)w_{it}^{T} \tag{3.1}$$

where P and T denote the permanent and transitory wage component respectively, w^P and w^T represent the "core" of each wage component, while g( ) and h( ) are functions capturing time shifts in the covariance structure. As in Chapter 2, a specification of the core permanent wage which will be adopted here is the random growth (RG henceforth) model, which allows each observational unit to have its own wage profile and whose EWMD estimation focuses on variances and covariances of intercepts and slopes of such individual profiles over the sample:

$$\begin{aligned} & w_{it}^{P} = \mu_i + \gamma_i t \\ & \begin{pmatrix} \mu_i \\ \gamma_i \end{pmatrix} \sim \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_\mu^2 & \sigma_{\mu\gamma} \\ & \sigma_\gamma^2 \end{pmatrix} \right] \end{aligned} \qquad (3.2).$$ [35]

An alternative specification which will also be analysed in this chapter allows permanent wages to follow a random walk process (RW henceforth);

$$\begin{aligned} & w_{it}^{P} = w_{it-1}^{P} + \xi_{it} \\ & w_{i0}^{P} = \mu_i \\ & \mu_i \sim (0, \sigma_\mu^2) \\ & \xi_{it} \sim WN(0, \sigma_\xi^2) \end{aligned} \qquad (3.3)$$

In the RG case permanent wages are supposed to evolve along linear profiles whose second moments have implications for the theory behind observed wage dynamics. The RW model, on the other hand, is more of a purely statistical kind, and is aimed at capturing the high level of wage persistence through the unit root hypothesis. As stressed in Baker [1997], such an outcome could arise from low rates of human capital depreciation or the impact of macroeconomic conditions via implicit contracts.

---

[35] Theoretical underpinnings of the model and their implications in terms of restrictions on the sign of $\sigma_{\mu\gamma}$ are discussed in Chapters 1 and 2.

The two models generate differing restrictions on the structure of second moments, in particular, while the RG imposes a quadratic dependence of covariance elements on calendar time (see equation 2.3), the RW implies a linear trend:

$$E^{RW}(w_{it}^{P}w_{is}^{P}) = \sigma_{\mu}^{2} + \min(t,s)\sigma_{\xi}^{2} \tag{3.4}$$

Moreover, both the models allow for mobility of permanent wages, the RG one through the sign of the convergence parameter ($\sigma_{\mu\gamma}$, see the discussion in Chapter 2), the RW model via the size of the white noise variance σ_{ξ}^{2}, a larger value implying greater scope for permanent wages to be reshuffled (I borrow this expression from Baker and Solon [1998]) with respect to their lagged values.

As far as the transitory component is concerned, its specification will follow the previous Chapter by hypothesising an ARMA(1,1) in which the variance of initial conditions is modelled separately from the white noise variance (the two parameters are indicated by σ_{0}^{2} and σ_{ε}^{2}, respectively, in the tables reporting results, while ρ and θ correspond to the AR and MA parameter respectively). This specification allows transitory mobility to be analysed in that it yields estimates of transitory wage correlation.

As anticipated in the Section 3.1, a central and qualifying difference with respect to the models of Chapter 2 is given by the specification of the time varying loading factors. These parameters are meant to capture the effect of forces which inflate-deflate the distribution of the two wage components, but leaving their ranks unaltered, thus impacting on the relative importance of the two components (and hence on persistence), but not on the extent of mobility within each of them. In Chapter 2, such loadings were specified as cubic functions on calendar time; here we follow Dickens [1996] and specify them as flexible time shifters, i.e. each time period has its couple of shifters, one for each component:

$$w_{it} = \pi_t w_{it}^P + \tau_t w_{it}^T \tag{3.5}$$

where π_t and τ_t indicate the shifters on the permanent and transitory component respectively, with the parameters for the first period set to 1 for identification. This specification implies the following restriction on second moments:

$$\begin{aligned} E(w_{it}w_{is}) &= \left(\sum_{t=0}^{T-1} d_t\pi_t\right)\left(\sum_{s=0}^{S-1} d_s\pi_s\right)E(w_{it}^P w_{is}^P) + \\ &\quad \left(\sum_{t=0}^{T-1} d_t\tau_t\right)\left(\sum_{s=0}^{S-1} d_s\tau_s\right)E(w_{it}^T w_{is}^T) \\ d_j &= I(j=k) \qquad j=t,s \qquad k=0,\ldots,T-1 \\ \pi_0 &= \tau_0 = 1 \end{aligned} \tag{3.6}$$

where the d_js are dummy variables indexing the rows and columns of the covariance matrix. Following this route, changes in the relative importance of the two wage components over time can be assessed without relying on specific assumptions on the functional form of the loadings.

3.3 The data utilised

The data set utilised in this study is a panel of individual wages referring to the 1979-1995 interval which has been made available by the INPS. The target population is the same as for the previous Chapter, i.e. dependent workers from the private non-agricultural sector of the economy; the available sample is a 1% random drawing of all the workers registered in the INPS archive during the period examined and born between 1928 and 1970. The data set has been built by merging the information contained in a form which refers to the worker with other information concerning the firm. On the workers side, the available information consists of the gross yearly wage (inclusive of any over-time and extraordinary compensation), year of reference, year of birth, gender, occupation and number of weeks worked. Information on the firm refers to its INPS identification code, size, geographical location and five digits industry.

The data set constitutes an unbalanced panel covering roughly 100,000 wage histories. Similarly to the sample of Chapter 2, attrition problems can potentially arise from non-random movements into and out from the data, caused by the same reasons outlined in the previous Chapter. Again, no formal control for attrition has been implemented due to the lack of instruments. However, some attention will be paid to the consequences of using the unbalanced panel instead of a balanced sample.

For the purposes of this study, I select full-time male workers employed on a regular basis born between 1930 and 1970 inclusive. Moreover, in order to improve the convergence properties of the GMM estimator, I also exclude the top and bottom 5 observations from each tail of the cross-sectional distributions. This yields an unbalanced panel where the total number of wage histories is 70,002 and whose structure is reported in Table 3.1, where diagonal elements give the cross-sectional dimension of the data and extra-diagonal element are the number of observations used in the estimation of the corresponding element of the covariance matrix.

The cross-sectional composition of the data with respect to some workers' characteristics is reported in Table 3.2. As can be seen, the cohort structure of the data reflects the entry of younger cohorts into the labour market; cohort turnover thus attenuates the progressive ageing of the sample with calendar time. We can also note a movement away from manual occupations, which can either reflect occupational mobility of older cohorts and a higher propensity of younger cohorts to be employed in non-manual jobs. A slight shift away from larger firms can be also observed, while the industrial structure tends to stay constant over time.

In order to construct the wage covariance matrix, the logarithms of real weekly wages (1995 prices) have first been adjusted for year, age and cohort effects. This has been done by regressing the 17 pooled cross-sections on a set of cohort dummies fully interacted with a quadratic in age and year dummies. The three effects are meant to capture business-cycle, life-cycle and productivity growth effects respectively, and the pooled cross-sections approach enables separate identification of age and birth cohorts. It's worth recalling from Chapter 2 that the aim of this initial adjustment is to remove the influence of structural factors (such as earnings progressions with age) which generate inequality between groups and could drive the results as an effect of changes of these characteristics within the sample through time. Moreover, the control for birth cohorts is very important in the INPS data since it can capture fixed differences in education between cohorts, thus, at least partially, coping with the non-availability of education among explanatory variables.

3.4 Patterns of inequality and mobility

Before proceeding to the formal analysis of the covariance structure, an useful starting point for the empirical investigation is to consider some descriptive statistics of both the marginal and the joint (over time) wage distribution.

Figure 3.1 reports the evolution of various indicators of (residual) cross-sectional wage dispersion computed from the 17 waves of the INPS panel after removing the structural effects mentioned above. Trends in the figure reproduce to some extent those reported by the literature reviewed in Chapter 1, with an initial phase of dropping inequality which stops in the early 1980s and is followed by a marked widening of wage differentials until the end of the data. The graph illustrates how increasing wage dispersion characterises not only the central-final part of the 1980s, as shown in Chapter 2, but is also a feature of the first half of the 1990s. The graph in the top left corner (stdev) plots the standard deviation of adjusted logarithmic weekly wages and shows how the reduction in inequality which occurs from the start of the data until 1982 is substantially neutralised by 1987. The re-opening of differentials is far from smooth and presents peaks in 1983, '85, '88 and '94. It is worth stressing how the last two years of data, subsequent to the 1993 bargaining round which completely abolished automatic wage indexation, are characterised by a pronounced increase in dispersion. The remaining panels of the figure (from l90_10 to l90_50) report the logs of the ratios between percentiles of the wage distribution, which, compared to sdlww, are robust to the presence of outliers. The evidence from l95_5, l90_10 and l75_25 suggests that the peaks in the growth of inequality are in fact a consequence of decompression at the tails of the distribution, being smoothed away the more robust the measure considered. From these panels it is also evident that wage dispersion grew quite slowly during the years immediately subsequent to the 1982 u-turn, while since 1987 the growth of inequality seems to be faster. Taking into account inequality trends at the bottom and top halves of the wage distribution (I50_10 and I90_50, respectively) it can be observed how the variation in the rate of growth of wage dispersion characterising the data since the second half of the 1980s can probably be ascribed to dynamics at the top of the distribution; on the other hand, the strong growth of the last two years of the sample seems to come from the bottom half.

The evidence shown so far concerns the marginal wage distribution: as such it is not informative about the degree of individual persistence (immobility) within the distribution which accompanies the widening of differentials. A first insight into this issue may be gained by considering some indicators of wage correlation and mobility.

Figure 3.2 reports the autocorrelation function of adjusted wages by various starting years. We can see how the patterns reported tend to reproduce the negative exponential shape of the autocorrelation function of an AR process. However there are some departures from it and, in particular, there seems to be an increase in persistence around the second half of the 1980s, the period which has been shown above to be characterised by an acceleration in the growth of dispersion; conversely, persistence drops faster over the early 1990s. The figure also suggests that persistence is increasing during the sample period, i.e. the autocorrelation functions tend to shift upwards as we move to more recent starting years. This last remark is confirmed by Figure 3.3, which plots wage autocorrelations by fixed lags over the sample period. Reported trends are typically increasing over time, although with some cycling. It is interesting to note how the downturns of such cycles tend to be placed in correspondence of the peaks of the standard deviation (the one reported in Figure 3.1); as an example, 1988 and 94 tend to be a local minimum in all the panels in the first row of the graph, suggesting that the distribution is more volatile in those years. As we move towards higher order lags, autocorrelation profiles tend to become flatter, indicating a greater stability of persistence for wages further apart, the effect of serially correlated wage shocks having washed away.

As mentioned in the previous Chapter, the correlation coefficient measures persistence in absolute terms and can be usefully accompanied by measures of wage mobility based on quantile transition matrices which focus on changes in relative wage ranks. Figure 3.4 plots the value of the mobility ratio and of the average absolute jump computed on ventile transition matrices for wages one, five and ten years apart; the first measure counts the average (over ventiles) frequency of cases changing ventile during the transition examined, while the second is concerned with the average absolute difference between departure and arrival ventile conditional on having moved. The joint implementation of the two indices allows us to focus not only on the number of workers moving, but also on the width of the transitions made.[36] Estimates of the frequency indicator are in line with the outcomes from the analysis of the correlation coefficient. In the short term mobility is decreasing over the sample period, with the exception of the last two years of the data; taking longer run measures into consideration, the evidence of dropping transition frequency is less apparent, mobility profiles tending to be flatter. Taking the width indicator into account, we can notice first of all how, under this respect, mobility is less pronounced, the maximum value plotted in the left panel being (differently from the frequency measure) even less than 50% of the "perfect mobility" value. Also, we can see how in this case, dropping mobility is more evident for longer lags.

[36] As an upper benchmark for these numbers one could use the values corresponding to the case of stochastic independence of wages in the two periods, which are .95 for the mobility ratio and 7 for the average absolute jump.

Overall, Figure 3.4 points towards a dynamic of the wage distribution in which transitions have been declining over time and of limited scope.

The evidence provided thus far shows that the degree of persistence of individual positions within the distribution tended to increase over the period analysed. If coupled with the increasing cross-sectional dispersion, this fact suggests that widening wage differentials permanently affected individual wage profiles. In the next section I use formal models of the wage covariance structure to directly address this issue.

3.5 Variance components models of the wage covariance structure

This section reports the results obtained by fitting the covariance structure implied by the models of Section 3.2 to the empirical covariance matrix estimated from the INPS panel.

The empirical covariance matrix for the whole sample of full time male workers is reported in Table 3.3: diagonal and super-diagonal elements are wage variances and autocovariances (respectively), with asymptotic standard errors reported in parentheses, while, below the diagonal, the corresponding correlation coefficients (which are the ones plotted out in Figures 3.2 and 3.3) have been computed. The covariance matrix elements are all significant at conventional confidence levels. Casual inspection of super-diagonal elements reveals how, typically, the autocovariance function presents the negative exponential shape noted above for the autocorrelation function; similar evidence is reported by Dickens [1996] for the UK and Moffitt and Gottschalk [1993] for the US, and in both cases it has been argued that these patterns can be picked up by some form of autocorrelation in the transitory component. It can be observed how, compared to the analogous table in Chapter 2, here the evidence of historic time dependence in the covariance structure is less apparent, but still present (recall, for example, the concavity characterising the autocorrelations of figure 3.2 over the 85-88 interval). We can also observe how comparable points in the two cases (say, for example, the variance of 1979) are, numerically, quite different, figures being lower in the case of the present Chapter: this is due to differences in the first stage regression, in particular to the fact that in Chapter 2 cohort effects are not fully interacted with age and year effects, thus leaving more variation in the data.

3.5.1 Estimates for the whole sample

A first group of results for the whole sample is reported in Table 3.4, namely models with a RW or RG specification of the permanent component, while the transitory wage is assumed to be ARMA(1,1); in both cases, no time shifters are allowed, so that the whole dynamics of the covariance structure are picked up by the linear and quadratic terms in calendar time within the permanent wage and by the correlation coefficient within the transitory one.[37]

Parameter estimates are well determined and indicate a substantial growth of permanent wage dispersion in both cases; in particular this is in the order of 64% in the RW case and 55% for the RG specification.[38] Moreover, the RG model indicates that workers with a growth rate parameter one standard deviation above the mean will experience a 24% growth of permanent wages over the sample period. A second

---

[37] Specification of the RG model assumes that the time trend is common to all workers, i.e. $w_{it}^{p} = \mu_i + \gamma_i t$; experiments have also been made with individual specific time trends ($w_{it}^{p} = \mu_i + \gamma_i t_i$, with t_i measured as difference from t and the first year i is observed in sample) obtaining virtually the same results as the one presented.

[38] This figures are computed as $\{[(\sigma^2\xi(T-1)+ \sigma^2\mu)/ \sigma^2\mu]-1\}*100$ for the RW and as $\{[(\sigma^2\gamma(T-1)^2+2\sigma\mu\gamma(T-1)+ \sigma^2\mu)/ \sigma^2\mu]-1\}*100$ for the RG model.

relevant fact emerging from these estimates is the positive sign of the covariance between intercepts and slopes of the RG model ($\sigma_{\mu\gamma}$) which indicates divergence in permanent wage profiles over the life cycle. This supports the interpretation which was given to the negative estimate of $\sigma_{\mu\gamma}$ in the previous Chapter, i.e. that such an outcome was an effect of the wage indexation system: the data set of the current Chapter is less influenced by the compressionary effect of the *Scala Mobile*, so that the underlying tendency of diverging wage profiles is evident also between occupations. Taking the transitory wage into account, we can notice how, differently from Chapter 2, the ARMA specification is now supported by the data also in the presence of a dynamic permanent wage, a fact probably arising from the presence of higher variation in the data set of the present Chapter, which allows identification of more flexible specifications of individual wages. Estimates of the transitory wage parameters are fairly stable across the two models, the most relevant difference being in the MA parameter which, as an effect of the introduction of a quadratic term in the permanent covariance (i.e. for the RG specification), gets nearly halved in size. A final comment is deserved by the measures of fit. The sum of squared residuals is lower for the RG model, in line with what we would expect from the fact that this model has one additional parameter; however, this doesn't hold for the χ^2 statistic, which is lower for the less parametrised RW model, suggesting that this specification provides a better description of permanent wage dynamics.

Table 3.5 reports estimates from the same specifications of the permanent and transitory wage, but including period specific time shifters on the two wage components. By comparing these results with the ones from the previous table we can see how parameter estimates for both the transitory (an exception is σ_0^2) and the permanent wage get inflated by the inclusion of the loading factors. On the other hand, the loadings are smaller than one and, for the RG specification of the permanent wage, they are monotonically decreasing over time. However, this doesn't mean that permanent wage covariance is decreasing over time, given that we have to consider the whole impact of estimated parameters, i.e. both the loadings and the dynamics within the permanent wage components as captured by $\sigma_{\xi}^2, \sigma_{\gamma}^2$ and $\sigma_{\mu\gamma}$.

To better assess the dynamics of covariance components, Figure 3.5 plots out the Table's predictions in terms of variance decomposition. These predictions are obtained utilising parameter's estimates in the formulas given in (3.7) below, where $E^P(w_{it}w_{is})$ denotes the predicted permanent covariance structure and $E^T(w_{it}w_{is})$ is the predicted transitory covariance structure, while predicted total covariance results from the sum of the two components:

$$\begin{aligned}
&E^P(w_{it}w_{is}) = (\textstyle\sum_{t=0}^{T-1} d_t \pi_t)(\sum_{s=0}^{S-1} d_s \pi_s) E(w_{it}^P w_{is}^P) \\
&E^T(w_{it}w_{is}) = (\textstyle\sum_{t=0}^{T-1} d_t \tau_t)(\sum_{s=0}^{S-1} d_s \tau_s) E(w_{it}^T w_{is}^T) \\
&d_j = I(j = k) \qquad j = t, s \qquad k = 0, \ldots, T-1 \\
&\pi_0 = \tau_0 = 1
\end{aligned} \tag{3.7}$$

In both cases, it can be observed how permanent variance accounts for the largest share of the level of total variance in each time period and, in parallel to the growth in total variance, the permanent variance profile is increasing over the sample period, while the transitory variance is roughly constant. The estimates underlying the figure imply that in the RW case, total variance increased by 54% from 1982 to 1995, of which 90% can be ascribed to the permanent component; corresponding figures for the RG model are 50% and 77%. Estimates of the core permanent component in the two cases still point towards a context of highly persistent, if not diverging, wage dynamics, and it is probably the large and positive estimate of $\sigma_{\mu\gamma}$ in the RG case which imparts the decreasing pattern to the loading factors for the permanent component, counterbalanced by the increasing transitory dispersion in the initial years of the sample. Combining the estimates of σ_{γ}^{2} and of the loading factors, the RG model predicts a 28% growth of permanent wages over the 17 years period for workers one standard deviation above the mean in the distribution of γ. Again, the χ^2 statistic favours the RW permanent wage specification, despite having one parameter less than the other model. Considering this fact in conjunction with the similar conclusion drawn from the previous table, it seems that the data support a picture of individual wage dynamics characterised by the high persistence of random walk processes, rather than evolving according to theoretically derived linear profiles. In any case, both the RG and RW specifications lead us to rule out the possibility of wage convergence at this stage of the analysis.

Going back now to the dynamics of transitory wage variance depicted in Figure 3.5, we can observe how the peaks characterising total variance in 1988 and 1994 are in fact a consequence of wage volatility. Transitory wage variance is increasing in the last years of the data; in particular, from 1989 to 1995, wage volatility grows by 76% in the RW case and by 59% in the RG one. Thus, apart from the peak of 1994 which, similarly to the one which can be observed in 1988, could in part arise from the higher turmoil characterising the distribution in these years and which produced dispersion at the tails (as Figure 3.1 showed), increasing volatility can be detected within the distribution in the last part of the period under investigation, a fact which accords with the higher institutional flexibility which characterises the Italian labour market over the late 1980s and the 1990s. This is also in line with the fact that, as observed when commenting on Figure 3.4, frequency mobility measures tend to drop more evidently in the short than in the medium and long run over these years: wage immobility has more to do with shock persistence in this case and washes out after a few periods. By recalling the discussion of Section 1.4, this also means that while on the one hand cross-sectional inequality has a lower impact over individual life-cycles, on the other wage profiles are characterised by higher uncertainty, which could worsen workers' welfare.

As a next step in the analysis, Table 3.6 and Figure 3.6 report results obtained by restricting the attention to the balanced sample, i.e. by discarding those wage profiles for which observations are missing in any of the years considered. By doing so, it will be possible to get a feeling of the effects of panel attrition on the models under estimation. Moreover, the selected sample will correspond to a more homogeneous group, the cases being ruled out relating to workers moving into or out from private dependent employment or the labour force, in particular workers belonging to extreme birth cohorts beginning or ending their careers. Estimation results will thus be informative of the effects of (broadly defined) job stability on the covariance of wage components.

Compared with Table 3.5, reported parameter estimates present a smaller dispersion of initial permanent wages (σ_{μ}^{2}), which accords with the sample design; the growth parameters of the "core" permanent variance are also lower, and similar remarks apply to the parameters of the transitory wage. The covariance between intercepts and slopes of the RG model still indicates divergence of wage profiles over the life-cycle. On the other hand, the size of the loading factors rises for the permanent component, while for the transitory one this holds only over the first part of the sample period, a drop in the transitory loadings characterising the last part of the data. The fitting measures suggest a higher capability of the models in capturing empirical variation, a likely consequence of the higher homogeneity characterising this sub-sample.

Figure 3.6 shows the implication of these estimates in terms of variance decomposition. First of all we can notice how the peaks of total variance in 1988 and, especially, 1994 are smoothed out. By recalling that, for the whole sample, these were symptoms of wage volatility, the finding reflects the more stable nature of the sub-sample under investigation. A relevant difference with respect to the whole sample is given by the growth of estimated total variance, which, for the interval 1982-1995 amounts at approximately 90% for both specifications of the permanent wage. As we should expect from the sample design, permanent variance plays a predominant role in shaping overall dynamics, its contribution to total variance growth being 98 and 95% for the RG and RW model respectively; thus, a higher permanent wage homogeneity at the beginning didn't translate into higher homogeneity at the end of the period. Accordingly, we can observe how the last part of the period is characterised by a lower level of transitory variance with respect to the unbalanced case.

3.5.2 The covariates of permanent and transitory inequality

Results presented thus far show that although the aggregate divergence of wage differentials had an impact on both wage components, a major role was played by the dynamics of permanent wage variance. In what follows, advantage will be taken of the availability of observable workers and firms characteristics in the INPS archive in order identify the directions along which permanent inequality and wage volatility grew more markedly.

A first way by which this purpose can be pursued is to remove also the effect of observable characteristics from the first stage regression; this amounts at abstracting from the effects of between groups wage differentials and the resulting covariance decomposition will thus be within the cells defined by the additional controls adopted in the first stage regression. A comparison of such within-groups analysis with results from Section 5.1.1 will then be informative about the effects of between-groups differentials on permanent and transitory wage variation. Among the (few) observable characteristics available in the INPS panel, here I focus on occupation, industry and firm size. To control for them, I add dummy variables defined according to the splits reported in Table 3.2 in the first two cases and the logarithm of firm size in the third to the age, cohort and time effects. Results obtained are reported in Table 3.7 and Figure 3.7. Reported specifications of the permanent component are the RW model for the within occupations analysis and the RG one in the remaining cases. While for the sample with occupational effects removed the RG model didn't converge to a well determined vector of estimated parameters, in the remaining cases the RG specification of the permanent wage provided a much better fit to the data, with the $\chi 2$ approximately halving as a consequence of the additional parameter. Such evidence contrasts with results from the previous analysis, when it was argued in favour of the RW specification and suggests that, those conclusions were in part due to persistence in wage differentials between firms of different size and industrial sector.

Looking at the "core" permanent wage for models with RG specification (i.e. the last three columns of Table 3.7), we can notice how the covariance between intercepts and slopes of individual profiles is now negative, indicating convergence. The dynamics of permanent variance are, for what concerns the two central columns of the table, still growing, as signalled by the loading factors and confirmed by the graph: after removing the effects of these observables, the impact of the forces generating permanent inequality is transferred from the interplay between fixed and time varying heterogeneity to shifters which affect workers independently from their position in the distribution of permanent wages. This is true also for the last column (in which the three effects are removed simultaneously), but the level of initial permanent dispersion is (not surprisingly) roughly half that in the other cases, so that the permanent variance profile plotted out in the graph displays a less evident growth. Focusing on the first column of the table and comparing it to its counterpart of Table 3.5, we can observe for the within occupation covariance structure a lower level of RW initial permanent dispersion, which is in line with the fact that wages are net of occupational dummies, and a larger white noise variance (σ_{ξ}^2), which means that wage profiles present larger variability around their long-run level, an occurrence which is, again, in accordance with the fact that between occupation differences have been taken out from adjusted wages. The loading factors in the first column of Table 3.7 are smaller and monotonically decreasing through time, so that a larger σ_{ξ}^2 doesn't impact on the dynamics of overall permanent variance.

The panels of Figure 3.7 clearly suggest that between occupations effects are, among those considered, the most important in determining permanent and, consequently, total wage inequality: the within-occupations profile of total variance is basically flat all over the period, while similar conclusions may not be drawn when within firms size or industries (but between occupations) differentials are taken into account. Thus, between occupations wage differentials seem to play a role in determining the level of wage persistence; given that a worker's occupation is probably the best proxy for permanent skills available in the INPS data, such evidence points towards a role for skill biased changes in the labour market in explaining the widening wage distribution. Of course, this doesn't explain the nature of such changes, which may well come from market forces or from a shift in the priorities of wage bargainers after the era of wage egalitarianism.

Finally, taking wage volatility into account, the figure provides a profile of transitory variance which is similar across cases and also resembles the ones of Figure 3.5; in particular, the growth of transitory variance over the 1989-95 interval ranges from 28% in the within size sample to 38% in the within occupations one. Combining these findings with their counterparts emerging from Figure 3.5, the growth of wage instability characterising the last years of the data appears to spread both between and within the cells defined by observable workers characteristics.

Evidence from the within groups analysis shows the importance of occupational differentials for the dynamics of variance components; a further investigation of this finding (along the lines of what has been done in Chapter 2) is provided by the first two columns of Table 3.8, which refer to models estimated on sub-samples defined according to workers' occupations. In order to perform such an exercise, the current workers occupation has been re-coded into a time invariant equivalent one, namely the weighted average of annual classifications with weights given by annual weeks worked,[39] so that wage profiles are not "cut" in correspondence of steps in the occupational career.

For both sub-groups, RW specifications of the permanent wage are estimated. While for the blue collar sample the RG model yielded a value of the χ^2 statistic higher than the RW's one, for white collar workers it failed to converge, signalling that it tends to overparametrise the data. Considering parameter estimates it can be

[39] In order to cope with the arbitrariness of this criterion, a dummy signalling whether in a given year the worker belongs to an occupational group different from the time invariant equivalent has been included in first stage regressions.

seen how the "core" components of the permanent wage are fairly similar across sub-samples, with the estimate of σ^2_μ, which is less precise (but still highly significant) for white collar data; moreover, the white noise variance within the RW (which captures the linear growth of permanent variance through time) is also similar to the one for the whole sample in Table 3.5. The parameters for the "core" transitory wage are instead less similar, white collar workers presenting a higher degree of serial correlation, meaning that wage shocks are more persistent for this group. Thus, the two occupational groups seem to be characterised by similar degrees of permanent wage mobility, while a lower level of transitory mobility is evident in the white collar sample. A difference may also be observed for what concerns the behaviour of the loading factors; in particular, for the permanent components of the white collar sample they start being larger since the second half of the 1980s, while for the transitory wage larger loadings for white collars may be observed over the end of the 1980s decade and the early 1990s.

In order to gain some additional insights into the occupational differences in the wage covariance structure, I proceed by jointly modelling the data from the two groups. The fact that some parameters tend to be similar across groups points towards the opportunity of a more parsimonious joint modelling, which would also allow direct tests of the statistical significance of such similarities. With this aim the two occupational covariance vectors have been stacked into a single vector of empirical moments[40], on which restrictions have been imposed assuming that each of the parameters of interest results from the sum of two parts, a base component

[40] The fourth moments matrix for this problem is block diagonal with blocks given by the matrices of the two groups.

and a second one representing the shift of the parameter in correspondence of second moments coming from the white collar sample.

Attempts at estimating a "fully shifted" model, however, produced estimates of the core transitory wage which were opposite to the ones arising from the sub-sample analysis, in that the AR and MA parameter were smaller in absolute value for white collar workers, while some of the transitory loading factors implausibly turned out to be negative for this group [41], leaving the impression that simultaneously shifting the core transitory wage and its loadings imposes too much structure on the data, and that restrictions have to be put on the transitory wage specification.[42] A first attempt has been made by restricting the core of the transitory component to be the same across groups, while letting the loadings vary: this produced a value of the χ^2 statistic of 1611.22 which, as we will see below, is larger than the one obtained by following the alternative route, despite corresponding to a model with 12 additional parameters; resulting estimates of the transitory loadings for the white collar sample were in general much smaller than the ones in the second column of Table 3.8, in particular being between .05 and .15 from 1989 to 1995. Thus the alternative route was taken, so that the specification of the occupationally shifted covariance structure model restricts the transitory wage loading factors to be the same across occupations, while the core parameters are allowed to vary: we can interpret this restriction by assuming that the loadings pick-up the effect of economy-wide shocks which impact on volatility irrespective of occupation, while the ARMA

[41] Negative loadings imply that workers on one side (with respect to the mean) of the distribution of the core wage happen to be on the other side of the distribution of the overall wage as a result of the interaction with the loading factor, an outcome hardly making sense whatever the nature of the forces behind the loading factors on the transitory wage.

[42] Attempts were also made by using this general model with only some of the transitory loadings restricted to be equal across occupational groups, namely the 1983, 85 and 88 ones, the choice being dictated by comparison of the within groups estimates; results similar to the one reported for the fully shifted model arose also with this specification.

parameters capture occupation specific serially correlated transitory shocks. The "preferred" specification is thus:

$$
\begin{aligned}
w_{it} &= (\pi_t + \pi_t^{wc} d_i)(w_{it-1}^P + \xi_{it}) + \tau_t[(\rho + \rho^{wc} d_i)v_{it-1} + \varepsilon_{it} + (\theta + \theta^{wc} d_i)\varepsilon_{it-1}] \\
w_{i0}^P &\sim [0,(\sigma_\mu^2 + \sigma_\mu^{2wc} d_i)] \\
\xi_{it} &\sim [0,(\sigma_\xi^2 + \sigma_\xi^{2wc} d_i)] \\
v_{i0} &\sim [0,(\sigma_0^2 + \sigma_0^{2wc} d_i)] \\
\varepsilon_{it} &\sim [0,(\sigma_\varepsilon^2 + \sigma_\varepsilon^{2wc} d_i)]
\end{aligned}
\qquad (3.8)
$$

where the *wc* superscript denotes shifts of the parameters in correspondence of white collar wages, while d_i is a dummy indicating white collar workers.

Results from the shifted covariance model estimates are reported in the last column of Table 3.8. The core of the permanent wage resembles the estimates for the blue collar sub-sample, while occupational shifts of its parameters are not significant, mirroring what has been noticed above when considering the two groups separately and confirming that the process governing core permanent wages is similar across occupations. The variance of initial conditions of the transitory wage also presents a statistically insignificant occupational shift. On the other hand, the ARMA parameters present significant differences across occupations, and, in particular, suggest that transitory shocks in the white collar distribution are more concentrated and more persistent.

The behaviour of the loading factors for the permanent wage is quite interesting. During the first part of the period considered, estimates of these parameters are fairly similar across occupations, with some of the white collar shifters being significantly negative, although of moderate size. Since the second half of the 1980s, and especially over the 1990s, this tendency reverts, and the white collar sub-sample attracts a considerable positive difference in these parameters with respect to blue collar data. Matched with the evidence about the core permanent wage discussed above, this implies that differences in persistence across occupations can be ascribed to factors which inflate the distribution and affect all individuals within one sub-group in the same fashion, rather than to persistent heterogeneity across individuals within that group. For example, this could result from contract provisions which differentiate wage dynamics between the two sub-groups, but not among individuals within each group, and can also explain the acceleration of aggregate inequality growth at the top half of the distribution over this period.

Figure 3.8 plots out predicted variance decomposition from this last model[43] and shows how the dynamics of wage dispersion differ considerably across sub-samples, a fact which, given the above discussion, can be ascribed to the loading factors of the permanent wage and to the parameters of the ARMA transitory process. Blue collar data present a profile of total wage dispersion which is, on the whole, decreasing up until 1993, while the peak observed in 1994 for the whole sample is still evident, and it is, as in that case, picked up by the transitory component. Overall inequality dynamics are accounted for by the transitory wage over the last part of the period. The right panel of the graph shows how both total and permanent differentials steadily grew after 1982 for white collars; moreover, while permanent and transitory wage variance are at the same level during the early 1980s, transitory differentials remain roughly constant until the end of that decade, implying that persistence is growing. The increasing volatility of the early 1990s also characterises white collar data, by an extent which is amplified if compared with the

[43] Only estimated occupational shifters marked by asterisks in Table 3.8 (i.e. statistically significant at least at the 20% level) have been used to construct the graph.

sample of manual workers, in accordance to the differences in the ARMA parameters estimated above.

3.6 Summary and conclusions

This Chapter has looked at the permanent/transitory nature of male wage inequality dynamics using a large panel of INPS wage data over the 1979-1995 interval. Descriptive statistics of the marginal wage distribution have shown that after the 1982 u-turn already documented in Chapter 2, cross-sectional differentials kept on growing up until the end of the data, in particular for what concerns the late 1980s-early 1990s interval and the top half of the distribution. Indices of correlation and mobility have been used to show that these trends were paralleled by increasing persistence, especially in the short and medium terms.

Formal models of the wage covariance structure have been estimated by Equally Weighted Minimum Distance. Alternative specifications of the permanent wage have been estimated, namely random walk and random growth models, showing that the former tends to provide a better description of permanent wage dynamics. Moreover, both specifications' estimates show the absence of wage convergence over the life-cycle; given that the data set of the current Chapter covers a period less influenced by the Scala Mobile than the one used in Chapter 2, this reinforces the conclusion advanced there, when convergence was ascribed to the effects of the egalitarian wage indexation system.

Models with flexible time shifters on the two wage components have been estimated to assess changes in the covariance structure through time. Results indicate that the overall growth in inequality was to a large extent driven by permanent differentials, which account for 70 to 90% (according to the specification of the permanent wage) of total inequality change. A role has been detected also for volatility, which contributes to growing dispersion over the early 1990s, a fact which accords with the increased "flexibility" of the labour market characterising these years. Thus, the relative importance of permanent inequality decreased in recent years due to a growth of wage uncertainty, which we have seen in Section 1.4 to have the potential for inducing welfare worsening effects. This conclusion is weakened when analysis is restricted to the balanced sub-sample, permanent differentials becoming predominant.

The analysis has next turned to the effect of observable workers' characteristics on covariance structure analysis. This has first been done by estimating our models on wages purged of the effects of between groups differentials, groups being defined by occupation, industry and firm size. It has been shown that much of the dynamics of total and permanent differentials arise from differences between occupations, while the other differentials considered do not alter conclusions reached about the dynamics of permanent and total inequality. On the other hand, the increasing volatility characterising the last part of the period analysed still persists after jointly removing the effects of observables, meaning that it spreads through the wage distribution both between and within the cells defined by the controls adopted.

Occupational differences have then been analysed in greater detail. Sub-samples defined according to occupation have been constructed and a method for shifting covariance parameters according to a worker's occupation has been proposed. Results show that permanent wages evolved according to random walk processes whose parameter estimates are similar across occupations, while estimates of time shifters for these processes indicate the presence of positive differentials in favour of white collar workers since the second half of the 1980s. On the other hand, ARMA transitory wages are significantly more persistent for white collar workers, while the loadings of this wage component have been restricted to be the same over occupations. The permanent wage structure thus reflects the use of wage premia in favour of white collar workers which, as mentioned in Chapters 1 and 2, became more frequent during the 1980s. However this has not generated differences in the dynamics of wage profiles within the white collar permanent wage distribution, but has rather shifted it with respect to the one for manual workers.

3. Tables and graphs

Table 3.1: The panel structure

70002	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
1979	55679																
1980	54537	58484															
1981	54494	57543	60807														
1982	54463	57345	59855	62706													
1983	54684	57544	59937	61958	64581												
1984	55201	58094	60476	62368	64292	67181											
1985	55149	58042	60420	62297	64171	66812	67799										
1986	55069	57954	60326	62222	64099	66695	67454	68277									
1987	55011	57910	60287	62174	64046	66651	67345	67919	68561								
1988	54847	57737	60100	61983	63846	66454	67128	67677	68092	68628							
1989	52402	55256	57609	59473	61317	63915	64580	65125	65500	65844	66266						
1990	50423	53154	55478	57399	59132	61701	62353	62916	63247	63581	63527	64121					
1991	48005	50621	52782	54676	56367	58901	59531	60078	60402	60724	60649	60720	61305				
1992	45119	47649	49714	51417	53093	55580	56196	56697	57038	57338	57224	57225	57137	57926			
1993	40466	42859	44784	46294	47945	50377	50988	51459	51784	52072	51971	51942	51811	51689	52650		
1994	38677	41051	42982	44498	46134	48424	49027	49488	49806	50074	49957	49785	49580	49240	47884	50602	
1995	34685	36955	38808	40247	41809	43997	44552	45014	45307	45562	45453	45282	45067	44710	43378	44882	46071

Table 3.2: Sample descriptive statistics

	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
mean age	33.56	34.03	34.59	35.21	35.85	36.41	37.23	38.07	38.98	39.87	40.58	41.33	42.01	42.61	43.00	43.50	43.74
birth cohort 1930-34	0.11	0.11	0.10	0.10	0.10	0.09	0.09	0.09	0.09	0.09	0.08	0.07	0.06	0.05	0.04	0.02	0.02
birth cohort 1935-39	0.17	0.17	0.16	0.16	0.15	0.15	0.15	0.14	0.14	0.14	0.14	0.14	0.13	0.12	0.11	0.10	0.08
birth cohort 1940-44	0.18	0.17	0.17	0.16	0.16	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.16	0.16	0.15	0.15	0.13
birth cohort 1945-49	0.20	0.19	0.19	0.18	0.18	0.18	0.17	0.17	0.17	0.17	0.17	0.18	0.18	0.19	0.19	0.20	0.21
birth cohort 1950-54	0.17	0.17	0.17	0.16	0.16	0.16	0.16	0.15	0.15	0.15	0.15	0.16	0.16	0.17	0.17	0.18	0.19
birth cohort 1955-59	0.13	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.15	0.15	0.16	0.17
birth cohort 1960-64	0.04	0.06	0.07	0.09	0.10	0.11	0.11	0.11	0.11	0.11	0.12	0.12	0.12	0.12	0.13	0.13	0.14
birth cohort 1965-70	0.00	0.00	0.00	0.01	0.01	0.02	0.03	0.03	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.06
blue collar	0.72	0.71	0.71	0.71	0.70	0.71	0.71	0.70	0.70	0.69	0.69	0.68	0.67	0.67	0.66	0.65	0.65
white collar	0.27	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.29	0.29	0.30	0.31	0.31	0.31	0.32	0.32
manager	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.03
size<15	0.18	0.18	0.18	0.19	0.20	0.21	0.22	0.22	0.21	0.21	0.20	0.20	0.19	0.19	0.20	0.20	0.20
15<=size<100	0.25	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.28	0.28	0.28	0.28	0.27	0.27	0.28	0.28
100<=size<500	0.20	0.20	0.20	0.19	0.19	0.19	0.19	0.19	0.19	0.20	0.20	0.20	0.21	0.21	0.20	0.20	0.20
500<=size	0.37	0.36	0.35	0.35	0.34	0.33	0.33	0.32	0.32	0.32	0.32	0.32	0.33	0.33	0.32	0.32	0.32
stone,clay and glass; basic metal industries; mining	0.09	0.09	0.09	0.09	0.09	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.07	0.07	0.07
food, wood and paper	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.09	0.09	0.10	0.10	0.10
textiles	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04	0.04	0.04
fabricated metal products and machinery	0.28	0.28	0.28	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27	0.27
energy and chemicals	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.10	0.09	0.09
constructions	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.10	0.10
transports and communications	0.08	0.07	0.07	0.07	0.07	0.07	0.08	0.07	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08	0.08
insurance, banking and financial s.	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.07	0.08
retail trade and other services	0.09	0.09	0.09	0.09	0.09	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
others	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
nobs	55679	58484	60807	62706	64581	67181	67799	68277	68561	68628	66266	64121	61305	57926	52650	50602	46071

Table 3.3: The covariance matrix of adjusted wages

	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
1979	0.158 (0.0018)	0.128 (0.0015)	0.115 (0.0013)	0.108 (0.0013)	0.110 (0.0013)	0.108 (0.0012)	0.109 (0.0013)	0.111 (0.0013)	0.115 (0.0013)	0.115 (0.0013)	0.113 (0.0013)	0.115 (0.0014)	0.115 (0.0014)	0.115 (0.0014)	0.111 (0.0015)	0.112 (0.0015)	0.110 (0.0015)
1980	0.838	0.148 (0.0016)	0.123 (0.0013)	0.112 (0.0012)	0.114 (0.0012)	0.111 (0.0012)	0.113 (0.0012)	0.113 (0.0012)	0.117 (0.0013)	0.117 (0.0013)	0.115 (0.0013)	0.117 (0.0013)	0.117 (0.0014)	0.118 (0.0014)	0.113 (0.0014)	0.114 (0.0015)	0.113 (0.0015)
1981	0.782	0.862	0.138 (0.0015)	0.114 (0.0013)	0.115 (0.0013)	0.111 (0.0012)	0.111 (0.0012)	0.111 (0.0012)	0.114 (0.0012)	0.113 (0.0012)	0.111 (0.0012)	0.113 (0.0013)	0.113 (0.0013)	0.114 (0.0013)	0.109 (0.0014)	0.110 (0.0014)	0.109 (0.0014)
1982	0.748	0.805	0.852	0.131 (0.0015)	0.119 (0.0013)	0.113 (0.0012)	0.112 (0.0012)	0.110 (0.0012)	0.112 (0.0012)	0.112 (0.0012)	0.110 (0.0012)	0.111 (0.0012)	0.110 (0.0012)	0.111 (0.0013)	0.107 (0.0013)	0.108 (0.0013)	0.107 (0.0014)
1983	0.734	0.785	0.817	0.872	0.143 (0.0016)	0.124 (0.0013)	0.122 (0.0013)	0.119 (0.0013)	0.121 (0.0013)	0.121 (0.0013)	0.117 (0.0012)	0.118 (0.0013)	0.118 (0.0013)	0.119 (0.0013)	0.115 (0.0014)	0.116 (0.0014)	0.115 (0.0014)
1984	0.719	0.767	0.789	0.823	0.869	0.143 (0.0016)	0.126 (0.0013)	0.122 (0.0012)	0.123 (0.0012)	0.122 (0.0013)	0.118 (0.0012)	0.120 (0.0013)	0.119 (0.0013)	0.120 (0.0013)	0.116 (0.0014)	0.117 (0.0014)	0.116 (0.0014)
1985	0.708	0.755	0.772	0.798	0.826	0.857	0.151 (0.0018)	0.132 (0.0014)	0.131 (0.0013)	0.130 (0.0013)	0.125 (0.0013)	0.126 (0.0013)	0.126 (0.0013)	0.126 (0.0014)	0.122 (0.0014)	0.124 (0.0014)	0.123 (0.0015)
1986	0.723	0.762	0.774	0.792	0.817	0.839	0.880	0.148 (0.0015)	0.139 (0.0014)	0.136 (0.0014)	0.130 (0.0013)	0.131 (0.0014)	0.130 (0.0014)	0.131 (0.0014)	0.127 (0.0014)	0.128 (0.0015)	0.128 (0.0015)
1987	0.720	0.760	0.763	0.771	0.795	0.811	0.838	0.896	0.161 (0.0017)	0.146 (0.0015)	0.139 (0.0014)	0.140 (0.0014)	0.139 (0.0014)	0.141 (0.0015)	0.135 (0.0015)	0.137 (0.0015)	0.137 (0.0016)
1988	0.700	0.736	0.741	0.750	0.772	0.782	0.808	0.855	0.882	0.171 (0.0019)	0.146 (0.0014)	0.146 (0.0014)	0.145 (0.0015)	0.145 (0.0015)	0.141 (0.0015)	0.142 (0.0016)	0.142 (0.0016)
1989	0.706	0.742	0.742	0.750	0.765	0.776	0.794	0.838	0.859	0.876	0.163 (0.0016)	0.151 (0.0015)	0.149 (0.0015)	0.149 (0.0015)	0.144 (0.0016)	0.146 (0.0016)	0.145 (0.0016)
1990	0.695	0.733	0.734	0.735	0.751	0.761	0.779	0.818	0.838	0.846	0.896	0.174 (0.0019)	0.157 (0.0016)	0.156 (0.0016)	0.152 (0.0016)	0.153 (0.0017)	0.153 (0.0017)
1991	0.685	0.721	0.724	0.725	0.742	0.750	0.766	0.803	0.824	0.832	0.875	0.897	0.177 (0.0017)	0.164 (0.0016)	0.157 (0.0017)	0.158 (0.0017)	0.159 (0.0018)
1992	0.671	0.708	0.711	0.710	0.729	0.736	0.750	0.788	0.811	0.814	0.853	0.869	0.900	0.186 (0.0018)	0.166 (0.0017)	0.165 (0.0017)	0.166 (0.0018)
1993	0.648	0.679	0.682	0.685	0.703	0.711	0.725	0.760	0.780	0.787	0.823	0.841	0.863	0.888	0.187 (0.0019)	0.170 (0.0018)	0.170 (0.0019)
1994	0.618	0.653	0.653	0.657	0.675	0.682	0.699	0.732	0.750	0.754	0.792	0.806	0.823	0.839	0.863	0.208 (0.0023)	0.178 (0.0019)
1995	0.612	0.651	0.651	0.654	0.675	0.679	0.700	0.733	0.756	0.757	0.793	0.813	0.833	0.852	0.865	0.864	0.205 (0.0021)

Note: asymptotic standard errors in parentheses, correlation coefficients below the diagonal

Table 3.4: Estimates for the whole sample without time shifters

	Random walk+ARMA(1,1)		Random growth+ARMA(1,1)	
$\sigma^2\mu$	0.1036	(0.0011)	0.1042	(0.0016)
$\sigma^2\gamma$			0.0002	(9.14E-06)
$\sigma^2\xi$; $\sigma\mu\gamma$	0.0042	(0.0001)	0.0002	(0.0001)
$\sigma^2\varepsilon$	0.0228	(0.0004)	0.0273	(0.0003)
σ^2_0	0.0559	(0.0014)	0.0484	(0.0017)
ρ	0.5362	(0.0170)	0.5859	(0.0135)
θ	-0.3194	(0.0241)	-0.1647	(0.0150)
fit	0.0040	2701.48	0.0033	2916.13
n.obs	70002		70002	

Notes: Asymptotic robust standard errors in parentheses; fit measures are the SSR (left) and the SSR weighted by the inverse of the fourth moments matrix (right). Columns are labelled according to the model specified for w_{it} (as defined in the text). Estimates shown refer to (in descending order): covariance structure of the core permanent wage (from $\sigma^2\mu$ to $\sigma^2\xi$; $\sigma\mu\gamma$), covariance structure of the core transitory wage (from $\sigma^2\varepsilon$ to θ).

Table 3.5: Estimates for the whole sample.

	π_t*Random walk+τ_t*ARMA(1,1)		π_t*Random growth+ τ_t*ARMA(1,1)	
$\sigma^2\mu$	0.1254	(0.0038)	0.1375	(0.0038)
$\sigma^2\gamma$			0.0013	(0.0002)
$\sigma^2\xi$; $\sigma\mu\gamma$	0.0051	(0.0007)	0.0067	(0.0011)
$\sigma^2\varepsilon$	0.0459	(0.0034)	0.0502	(0.0038)
σ^2_0	0.0323	(0.0035)	0.0200	(0.0034)
ρ	0.7323	(0.0219)	0.8430	(0.0086)
θ	-0.4319	(0.0133)	-0.4241	(0.0074)
$\pi80$	0.9739	(0.0063)	0.8968	(0.0112)
$\pi81$	0.9061	(0.0086)	0.7833	(0.0156)
$\pi82$	0.8562	(0.0104)	0.6986	(0.0187)
$\pi83$	0.8821	(0.0132)	0.6852	(0.0224)
$\pi84$	0.8585	(0.0146)	0.6363	(0.0242)
$\pi85$	0.8733	(0.0167)	0.6207	(0.0266)
$\pi86$	0.8782	(0.0188)	0.6035	(0.0286)
$\pi87$	0.9054	(0.0209)	0.6015	(0.0308)
$\pi88$	0.9043	(0.0218)	0.5783	(0.0317)
$\pi89$	0.8933	(0.0234)	0.5558	(0.0321)
$\pi90$	0.9078	(0.0249)	0.5467	(0.0331)
$\pi91$	0.9076	(0.0260)	0.5292	(0.0333)
$\pi92$	0.9155	(0.0270)	0.5166	(0.0340)
$\pi93$	0.8875	(0.0266)	0.4829	(0.0331)
$\pi94$	0.8967	(0.0270)	0.4711	(0.0335)
$\pi95$	0.9013	(0.0282)	0.4608	(0.0338)
$\tau80$	0.6490	(0.0428)	0.6480	(0.0335)
$\tau81$	0.6862	(0.0395)	0.7240	(0.0337)
$\tau82$	0.7174	(0.0367)	0.7598	(0.0323)
$\tau83$	0.7401	(0.0374)	0.7868	(0.0328)
$\tau84$	0.7636	(0.0363)	0.7982	(0.0323)
$\tau85$	0.7746	(0.0372)	0.8104	(0.0328)
$\tau86$	0.6729	(0.0328)	0.7389	(0.0287)
$\tau87$	0.6717	(0.0333)	0.7390	(0.0288)
$\tau88$	0.7406	(0.0360)	0.7851	(0.0314)
$\tau89$	0.6258	(0.0316)	0.6988	(0.0277)
$\tau90$	0.6495	(0.0357)	0.7174	(0.0303)
$\tau91$	0.6477	(0.0325)	0.7207	(0.0292)
$\tau92$	0.6782	(0.0337)	0.7481	(0.0304)
$\tau93$	0.7667	(0.0367)	0.8094	(0.0332)
$\tau94$	0.9053	(0.0414)	0.9147	(0.0380)
$\tau95$	0.8265	(0.0404)	0.8739	(0.0364)
fit	0.00016	911.19	0.00013	1020.24
n obs	70002		70002	

Notes: **Asymptotic robust standard errors in parentheses; fit measures are the SSR (left) and the SSR weighted by the inverse of the fourth moments matrix (right). Columns are labelled according to the model specified for w_{it} (as defined in the text). Estimates shown refer to (in descending order): covariance structure of the core permanent wage (from $\sigma^2\mu$ to $\sigma^2\xi$; $\sigma\mu\gamma$), covariance structure of the core transitory wage (from $\sigma^2\varepsilon$ to θ), loadings of the core permanent (the πs) and transitory (the τs) wage**.

Table 3.6: Estimates for the balanced sample.

	π_t*Random walk+τ_t*ARMA(1,1)		π_t*Random growth+ τ_t*ARMA(1,1)	
$\sigma^2\mu$	0.0782	(0.0023)	0.0779	(0.0037)
$\sigma^2\gamma$			0.0002	(0.0002)
$\sigma^2\xi$; $\sigma\mu\gamma$	0.0016	(0.0002)	0.0021	(0.0018)
$\sigma^2\varepsilon$	0.0279	(0.0017)	0.0333	(0.0026)
σ^2_0	0.0446	(0.0022)	0.0450	(0.0037)
ρ	0.7441	(0.0158)	0.7603	(0.0096)
θ	-0.3194	(0.0131)	-0.2600	(0.0105)
$\pi 80$	1.0365	(0.0058)	1.0078	(0.0308)
$\pi 81$	1.0076	(0.0075)	0.9542	(0.0553)
$\pi 82$	0.9739	(0.0089)	0.8985	(0.0748)
$\pi 83$	1.0418	(0.0111)	0.9387	(0.0999)
$\pi 84$	1.0450	(0.0130)	0.9204	(0.1175)
$\pi 85$	1.0883	(0.0151)	0.9370	(0.1380)
$\pi 86$	1.1167	(0.0176)	0.9421	(0.1558)
$\pi 87$	1.1773	(0.0203)	0.9731	(0.1772)
$\pi 88$	1.1892	(0.0221)	0.9630	(0.1904)
$\pi 89$	1.2038	(0.0241)	0.9554	(0.2028)
$\pi 90$	1.2377	(0.0258)	0.9622	(0.2174)
$\pi 91$	1.2533	(0.0274)	0.9543	(0.2281)
$\pi 92$	1.2837	(0.0291)	0.9565	(0.2403)
$\pi 93$	1.2883	(0.0300)	0.9391	(0.2470)
$\pi 94$	1.2997	(0.0306)	0.9262	(0.2541)
$\pi 95$	1.3533	(0.0323)	0.9432	(0.2688)
$\tau 80$	0.8338	(0.0210)	0.8001	(0.0326)
$\tau 81$	0.8275	(0.0254)	0.7818	(0.0335)
$\tau 82$	0.8516	(0.0279)	0.7972	(0.0317)
$\tau 83$	0.8761	(0.0306)	0.8215	(0.0339)
$\tau 84$	0.8354	(0.0305)	0.7890	(0.0335)
$\tau 85$	0.8629	(0.0362)	0.8175	(0.0375)
$\tau 86$	0.7090	(0.0276)	0.7024	(0.0297)
$\tau 87$	0.6692	(0.0288)	0.6789	(0.0299)
$\tau 88$	0.6344	(0.0327)	0.6514	(0.0323)
$\tau 89$	0.5690	(0.0224)	0.6013	(0.0255)
$\tau 90$	0.6083	(0.0320)	0.6276	(0.0330)
$\tau 91$	0.5653	(0.0218)	0.5966	(0.0265)
$\tau 92$	0.6000	(0.0263)	0.6266	(0.0296)
$\tau 93$	0.5903	(0.0271)	0.6244	(0.0286)
$\tau 94$	0.6982	(0.0296)	0.7088	(0.0317)
$\tau 95$	0.7475	(0.0321)	0.7606	(0.0326)
fit	0.00005	551.50	0.00004	544.76
n obs	29235		29235	

Notes Asymptotic robust standard errors in parentheses; fit measures are the SSR (left) and the SSR weighted by the inverse of the fourth moments matrix (right). Columns are labelled according to the model specified for w_{it} (as defined in the text). Estimates shown refer to (in descending order): covariance structure of the core permanent wage (from $\sigma^2\mu$ to $\sigma^2\xi$; $\sigma\mu\gamma$), covariance structure of the core transitory wage (from $\sigma^2\varepsilon$ to θ), loadings of the core permanent (the πs) and transitory (the τs) wage.

Table 3.7: Estimates for the whole sample adding the indicated controls to the first stage regression.

	Occupation		Firm's size		Sectoral affiliation		Occup., size &sect.	
$\sigma^2\mu$	0.0955	(0.0039)	0.0854	(0.0026)	0.0763	(0.0021)	0.0394	(0.0020)
$\sigma^2\gamma$			0.0001	(0.00001)	0.0001	(0.00001)	0.0001	(0.00001)
$\sigma^2\xi$; $\sigma\mu\gamma$	0.0122	(0.0013)	-0.0021	(0.0004)	-0.0022	(0.0002)	-0.0013	(0.0001)
$\sigma^2\varepsilon$	0.0396	(0.0026)	0.0400	(0.0020)	0.0413	(0.0019)	0.0418	(0.0018)
σ^2_0	0.0234	(0.0038)	0.0493	(0.0025)	0.0558	(0.0019)	0.0583	(0.0020)
ρ	0.7969	(0.0135)	0.7742	(0.0092)	0.7364	(0.0091)	0.7511	(0.0073)
θ	-0.4430	(0.0088)	-0.3575	(0.0083)	-0.3223	(0.0086)	-0.3193	(0.0071)
$\pi80$	0.8895	(0.0111)	1.0264	(0.0104)	1.0104	(0.0085)	1.0207	(0.0131)
$\pi81$	0.7704	(0.0115)	1.0124	(0.0179)	1.0251	(0.0136)	1.0411	(0.0203)
$\pi82$	0.6758	(0.0114)	1.0101	(0.0260)	1.0435	(0.0195)	1.0357	(0.0278)
$\pi83$	0.6693	(0.0136)	1.0816	(0.0370)	1.1082	(0.0272)	1.0967	(0.0376)
$\pi84$	0.6223	(0.0140)	1.0948	(0.0466)	1.1451	(0.0349)	1.1093	(0.0461)
$\pi85$	0.6043	(0.0150)	1.1736	(0.0602)	1.2342	(0.0453)	1.1951	(0.0587)
$\pi86$	0.5934	(0.0167)	1.2228	(0.0736)	1.3188	(0.0568)	1.2559	(0.0712)
$\pi87$	0.5914	(0.0178)	1.3074	(0.0903)	1.3954	(0.0691)	1.3019	(0.0837)
$\pi88$	0.5718	(0.0176)	1.3532	(0.1060)	1.4824	(0.0837)	1.3534	(0.0981)
$\pi89$	0.5539	(0.0182)	1.3825	(0.1213)	1.5375	(0.0979)	1.3831	(0.1101)
$\pi90$	0.5437	(0.0183)	1.4622	(0.1423)	1.6231	(0.1153)	1.4337	(0.1252)
$\pi91$	0.5361	(0.0184)	1.5133	(0.1618)	1.6850	(0.1321)	1.4930	(0.1423)
$\pi92$	0.5299	(0.0186)	1.5824	(0.1846)	1.7402	(0.1494)	1.5334	(0.1592)
$\pi93$	0.4973	(0.0175)	1.5957	(0.2028)	1.7650	(0.1664)	1.5116	(0.1716)
$\pi94$	0.5003	(0.0177)	1.6344	(0.2275)	1.8217	(0.1896)	1.5479	(0.1930)
$\pi95$	0.5090	(0.0190)	1.6756	(0.2559)	1.8608	(0.2146)	1.6077	(0.2201)
$\tau80$	0.6857	(0.0366)	0.7927	(0.0214)	0.8087	(0.0174)	0.8271	(0.0168)
$\tau81$	0.7809	(0.0361)	0.8177	(0.0236)	0.8328	(0.0209)	0.8487	(0.0200)
$\tau82$	0.8537	(0.0343)	0.8486	(0.0233)	0.8484	(0.0210)	0.8879	(0.0200)
$\tau83$	0.8574	(0.0348)	0.8754	(0.0244)	0.8643	(0.0219)	0.8991	(0.0210)
$\tau84$	0.8787	(0.0343)	0.8869	(0.0244)	0.8765	(0.0224)	0.9027	(0.0213)
$\tau85$	0.8908	(0.0353)	0.9066	(0.0255)	0.8873	(0.0237)	0.9150	(0.0227)
$\tau86$	0.7865	(0.0299)	0.8230	(0.0213)	0.7994	(0.0194)	0.8280	(0.0189)
$\tau87$	0.7836	(0.0309)	0.8191	(0.0223)	0.7963	(0.0207)	0.8219	(0.0200)
$\tau88$	0.8543	(0.0339)	0.8732	(0.0251)	0.8507	(0.0231)	0.8756	(0.0223)
$\tau89$	0.7307	(0.0286)	0.7814	(0.0213)	0.7518	(0.0191)	0.7716	(0.0183)
$\tau90$	0.7697	(0.0332)	0.8067	(0.0257)	0.7821	(0.0238)	0.8011	(0.0228)
$\tau91$	0.7623	(0.0304)	0.8006	(0.0232)	0.7807	(0.0209)	0.7902	(0.0198)
$\tau92$	0.7782	(0.0310)	0.8090	(0.0245)	0.8074	(0.0223)	0.7902	(0.0202)
$\tau93$	0.8369	(0.0328)	0.8438	(0.0265)	0.8345	(0.0242)	0.8111	(0.0215)
$\tau94$	0.9648	(0.0364)	0.9736	(0.0299)	0.9521	(0.0275)	0.9220	(0.0251)
$\tau95$	0.8612	(0.0332)	0.8851	(0.0270)	0.8627	(0.0255)	0.8066	(0.0214)
fit	0.0001	1009.80	0.0001	814.06	0.0001	745.77	0.0001	864.48
n obs	70002		70002		70002		70002	

Notes: Asymptotic robust standard errors in parentheses; fit measures are the SSR (left) and the SSR weighted by the inverse of the fourth moments matrix (right). Columns are labelled according to the additional controls adopted in the first stage regression. The core permanent wage is specified as a random walk in the "Occupation" column and as a random growth in the remaining columns. The core transitory wage is specified as an ARMA(1,1) in each column. Estimates shown refer to (in descending order): covariance structure of the core permanent wage (from $\sigma^2\mu$ to $\sigma^2\xi$; $\sigma\mu\gamma$), covariance structure of the core transitory wage (from $\sigma^2\varepsilon$ to θ), loadings of the core permanent (the πs) and transitory (the τs) wage.

Table 3.8: Estimates by time invariant equivalent occupational classification.

	Blue Collar Workers		White Collar Workers		Joint model with occupational shifters			
					Base parameter		Shifter	
$\sigma^2\mu$	0.0776	(0.0039)	0.0699	(0.0114)	0.0764	(0.0034)	-0.0013	(0.0069)
$\sigma^2\xi$	0.0053	(0.0010)	0.0043	(0.0018)	0.0052	(0.0009)	0.0009	(0.0014)
$\sigma^2\varepsilon$	0.0482	(0.0038)	0.0289	(0.0031)	0.0380	(0.0030)	-0.0123*	(0.0016)
σ^2_0	0.0392	(0.0037)	0.0443	(0.0117)	0.0393	(0.0033)	0.0019	(0.0071)
ρ	0.6368	(0.0198)	0.9390	(0.0120)	0.5861	(0.0222)	0.3397*	(0.0226)
θ	-0.3031	(0.0135)	-0.4655	(0.0301)	-0.2552	(0.0180)	-0.2351*	(0.0273)
$\pi 80$	0.9354	(0.0116)	0.9536	(0.0236)	0.9366	(0.0108)	-0.0086	(0.0255)
$\pi 81$	0.8487	(0.0162)	0.8529	(0.0268)	0.8536	(0.0146)	-0.0283	(0.0306)
$\pi 82$	0.7916	(0.0195)	0.7492	(0.0283)	0.8050	(0.0178)	-0.0881*	(0.0344)
$\pi 83$	0.7865	(0.0233)	0.7946	(0.0290)	0.8067	(0.0214)	-0.0635**	(0.0387)
$\pi 84$	0.7538	(0.0252)	0.7133	(0.0345)	0.7616	(0.0225)	-0.0549***	(0.0408)
$\pi 85$	0.7350	(0.0269)	0.7748	(0.0335)	0.7552	(0.0249)	-0.0324	(0.0434)
$\pi 86$	0.7260	(0.0292)	0.7887	(0.0389)	0.7365	(0.0264)	0.0212	(0.0451)
$\pi 87$	0.7116	(0.0300)	0.8717	(0.0558)	0.7250	(0.0273)	0.0767***	(0.0483)
$\pi 88$	0.7066	(0.0312)	0.8196	(0.0480)	0.7198	(0.0285)	0.0510	(0.0494)
$\pi 89$	0.6742	(0.0316)	0.8172	(0.0503)	0.6804	(0.0282)	0.1346*	(0.0519)
$\pi 90$	0.6573	(0.0315)	0.8405	(0.0552)	0.6655	(0.0283)	0.1559*	(0.0537)
$\pi 91$	0.6509	(0.0322)	0.8408	(0.0568)	0.6574	(0.0287)	0.1725*	(0.0549)
$\pi 92$	0.6281	(0.0317)	0.8526	(0.0585)	0.6327	(0.0281)	0.2198*	(0.0565)
$\pi 93$	0.6052	(0.0313)	0.7860	(0.0506)	0.6075	(0.0277)	0.1865*	(0.0530)
$\pi 94$	0.6011	(0.0312)	0.7813	(0.0509)	0.6062	(0.0281)	0.1605*	(0.0530)
$\pi 95$	0.6005	(0.0324)	0.8158	(0.0607)	0.6038	(0.0288)	0.2037*	(0.0557)
$\tau 80$	0.7130	(0.0415)	0.8344	(0.0556)	0.7589	(0.0360)		
$\tau 81$	0.7384	(0.0413)	0.8615	(0.0529)	0.8176	(0.0390)		
$\tau 82$	0.7721	(0.0381)	0.8396	(0.0476)	0.8361	(0.0380)		
$\tau 83$	0.8146	(0.0394)	0.8183	(0.0510)	0.8600	(0.0398)		
$\tau 84$	0.7849	(0.0383)	0.9265	(0.0501)	0.9044	(0.0408)		
$\tau 85$	0.8383	(0.0406)	0.8260	(0.0508)	0.8926	(0.0406)		
$\tau 86$	0.7147	(0.0347)	0.7964	(0.0502)	0.8063	(0.0376)		
$\tau 87$	0.7350	(0.0357)	0.7298	(0.0564)	0.8074	(0.0386)		
$\tau 88$	0.7850	(0.0387)	0.8090	(0.0532)	0.8662	(0.0401)		
$\tau 89$	0.6395	(0.0314)	0.8112	(0.0503)	0.7399	(0.0347)		
$\tau 90$	0.6912	(0.0376)	0.8133	(0.0539)	0.7814	(0.0395)		
$\tau 91$	0.6761	(0.0326)	0.8289	(0.0549)	0.7733	(0.0362)		
$\tau 92$	0.6893	(0.0330)	0.8801	(0.0558)	0.7982	(0.0375)		
$\tau 93$	0.7378	(0.0351)	0.9315	(0.0549)	0.8672	(0.0402)		
$\tau 94$	0.9095	(0.0421)	1.0092	(0.0584)	1.0402	(0.0459)		
$\tau 95$	0.7914	(0.0368)	0.9506	(0.0547)	0.9164	(0.0417)		
fit	0.000151	621.82	9.7E-05	607.48	0.0004	1387.24		
n obs	49011		19894		68905			

Notes: Asymptotic robust standard errors in parentheses; fit measures are the SSR (left) and the SSR weighted by the inverse of the fourth moments matrix (right). The first two columns are labelled according to the sub-sample used in estimation. Remaining columns are labelled according to the parameter component they refer to. The core permanent wage is specified as a random walk, the core transitory wage as an ARMA(1,1). Estimates shown refer to (in descending order): covariance structure of the core permanent wage (from $\sigma^2\mu$ to $\sigma^2\xi$), covariance structure of the core transitory wage (from $\sigma^2\varepsilon$ to θ), loadings of the core permanent (the πs) and transitory (the τs) wage. In the "Shifter" column *, ** and ***, denote significance at the 5, 10 and 20 percent levels respectively.

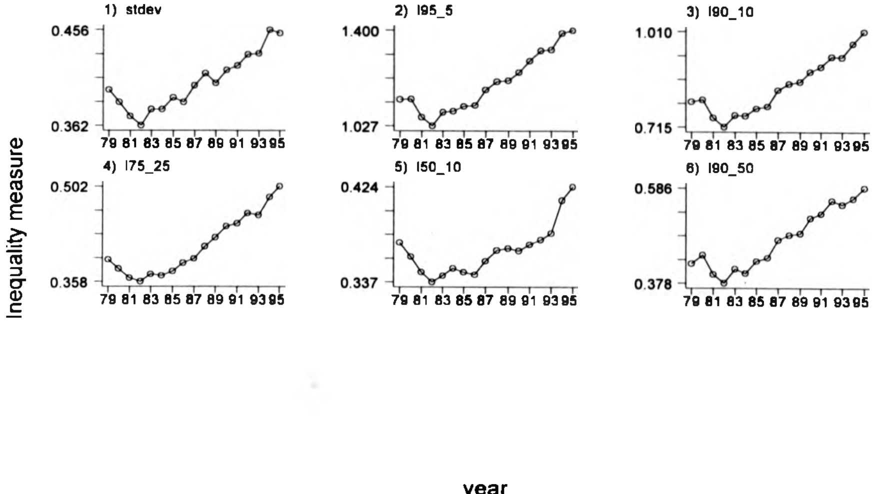

Fig. 3.1: Wage inequality dynamics

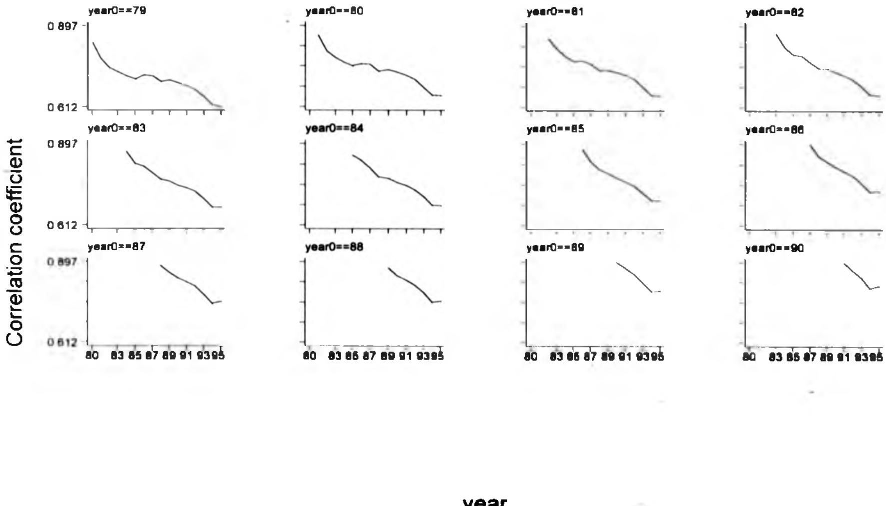

Fig. 3.2: Autocorrelations by starting year

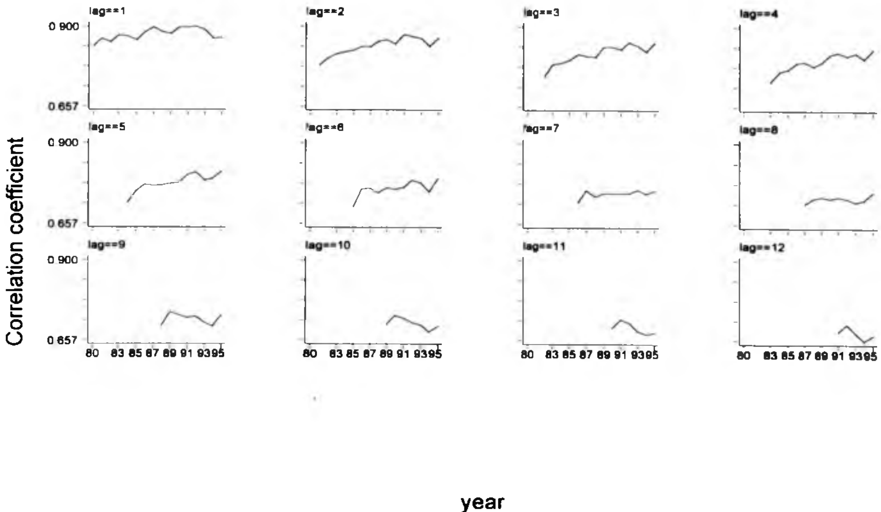

Fig. 3.3: Autocorrelations by lag

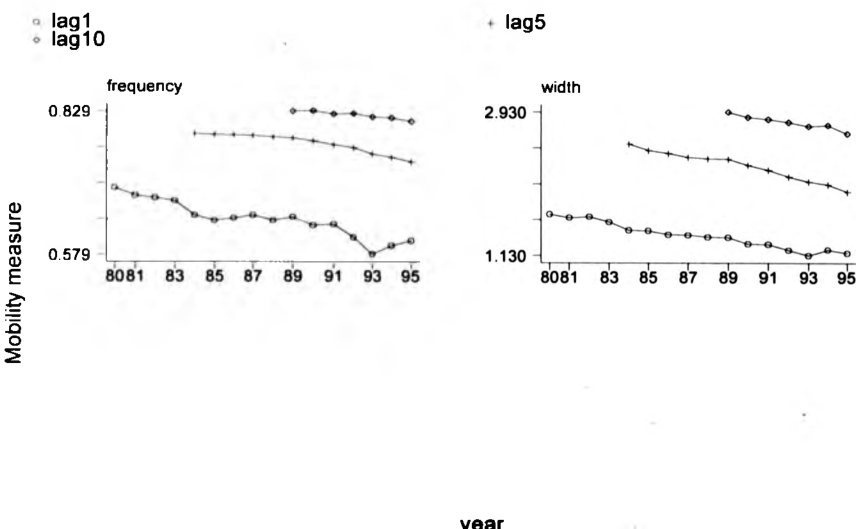

Fig. 3.4: Mobility indices by lag

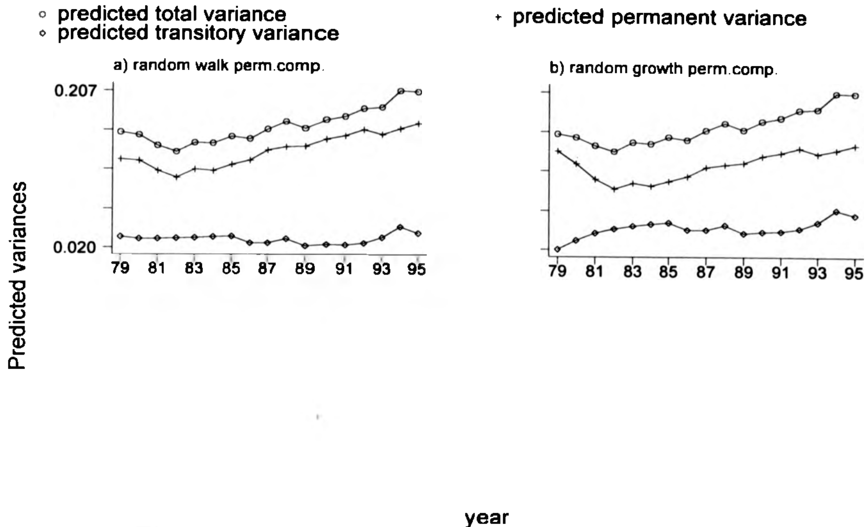

Fig. 3.5: Variance decomposition - Whole sample

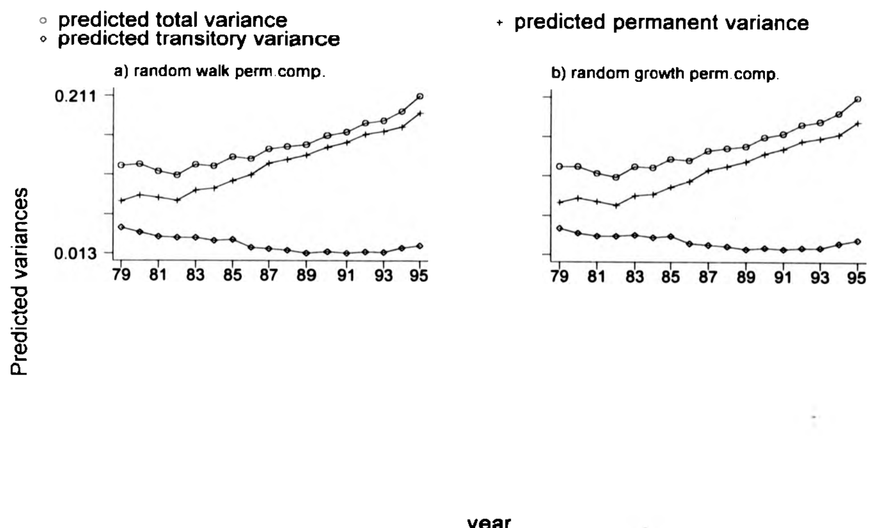

Fig. 3.6: Variance decomposition - Balanced sample

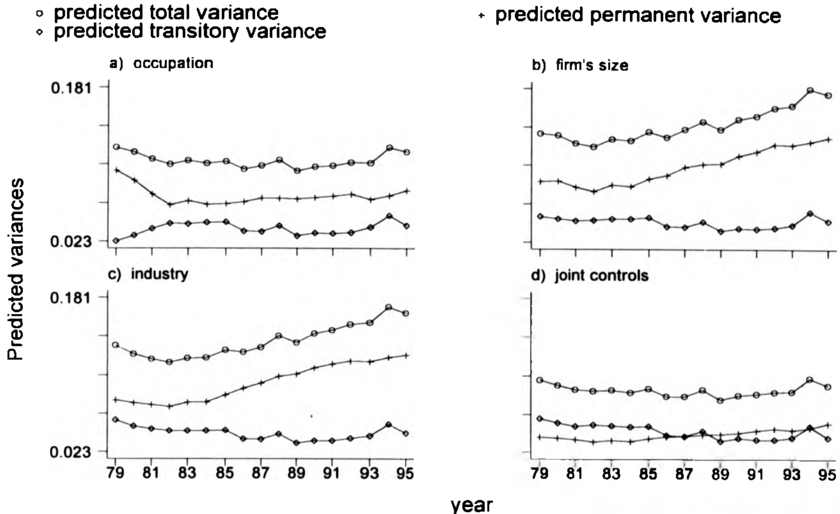

Fig. 3.7: Variance decomp.-Whole sample within groups

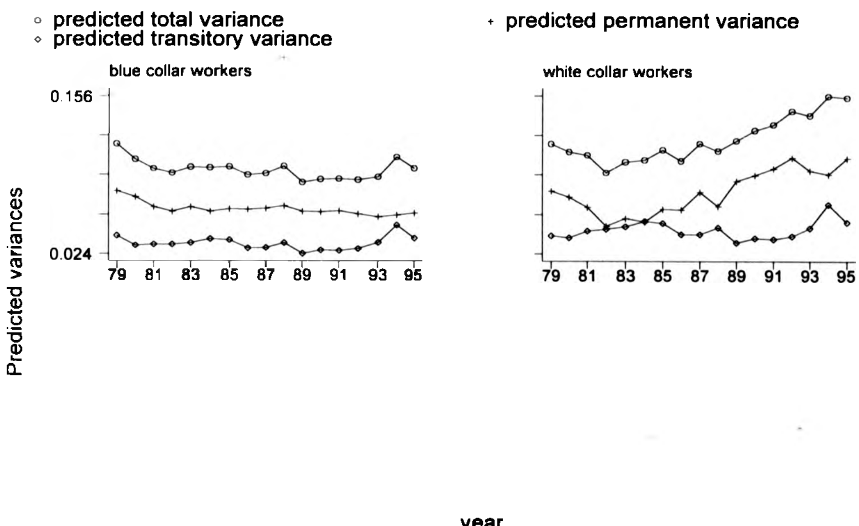

Fig. 3.8: Variance decomp. with occupational shifters

Chapter 4

Mobility at the bottom of the wage distribution

4.1 Introduction

This Chapter analyses individual transition probabilities at the bottom of the wage distribution using survey panel data from the Bank of Italy. As stressed in Chapter 1, increasing wage inequality involves a low-pay issue, with a growing proportion of the labour force paid below fixed “decency thresholds”, a fact whose policy relevance is witnessed in many countries by a renewed interest in minimum wage legislations. However, focusing on the incidence of low-pay at a point in time offers only an incomplete picture of the problem which has to be integrated with analyses of the degree of persistence of the low-pay status over individual careers. At one extreme, the bottom of the wage distribution could be characterised by a high degree of mobility, so that the experience of low-pay is shared among individuals over time. On the other hand, if low-pay is a persistent condition, workers are trapped in such “bad” jobs for a relevant portion of their career, so that the labour market produces inequality in a dynamic sense even if cross-sectional wage distributions are stable over time and the need for low-wage protection is more urgent. It is then important to analyse the degree of wage mobility across the low-pay threshold.

The analyses of this Chapter are focused on the econometric modelling of low-wage mobility, which can provide indications about the workers’ attributes relevant in generating mobility across the low-pay threshold. Moreover, such an exercise will allow investigation of the extent of true state dependence within aggregate low-pay persistence, thus shedding light on the effect of the past experience of low-pay in determining future low-pay episodes. Such an approach will then provide an alternative and complementary perspective on wage persistence with respect to the one of Chapters 2 and 3: while there the whole wage distribution was subsumed into a limited number of parameters and the effect of both observable and unobservable characteristics was assessed, the current Chapter places the emphasis on wage classes at the bottom of the distribution and on the effect of observable attributes in determining exits from and entries into them.

As pointed out in Section 1.6, various studies of wage mobility have been devoted to the econometric modelling of transition probabilities in recent years by treating the outcome mobile/not-mobile by means of discrete response models and conditioning it on a set of personal characteristics. However, as stressed by Bingley *et al.* [1995] and Stewart and Swaffield [1999] (S&S thereafter), the structure of the problem involves two potential sources of endogenous sample selection which, if not properly tackled, could bias parameter estimates in the transition probability equation. First of all, the analysis of transitions requires the conditioning of current wage states on their lagged values and as long as workers assignment to the initial conditions of the transition is correlated with unobservable workers characteristics, such a conditioning cannot be deemed exogenous. Secondly, mobility can only be analysed for workers with a valid wage at both ends of the transition and the presence of non-random exits from the wage distribution could induce a second source of endogenous selectivity. The modelling strategy of this Chapter will be focused on the initial conditions problem and, in particular, the model of S&S will be extended. Robustness of these results to the presence of attrition bias will be analysed in Chapter 5.

The Chapter is organised as follows. Section 4.2 describes the 1993 and 1995 waves of the SHIW data, which form the object of the analysis. Section 4.3 defines the low-pay thresholds and describes the characteristics of low-paid workers in terms of the *ceteris paribus* probability of being low-paid. Section 4.4 takes into account transitions out from and into the low-pay status: the econometric model of low-wage mobility is set out and results are presented. Section 4.5 analyses the impact of a more flexible specification of the transition equation, which accounts for the width of transition, on estimated parameters. Section 4.6 concludes, while a comparison between the data set used in this Chapter and the INPS data is reported in the Appendix.

4.2 The data

The data set utilised in this study is drawn from the 1993 and 1995 waves of the *Survey on Households Income and Wealth* (SHIW), a micro-data archive set up by the Bank of Italy with the aim of providing information on the economic behaviour of Italian households. Interviews have been conducted on an annual basis since 1977 and biannually from 1987 onwards. Although the sampling unit is the household, increasingly detailed information on labour market variables for individuals within the household has been made available in the recent waves of the survey[44].

The two waves utilised are the latest in the SHIW and various reasons dictated the choice. First of all, given that the focus of this study is the dynamic behaviour of wage earners and of their transitions within the wage distribution through time, the availability of a panel is crucial. A panel sub-section has been introduced in the SHIW data since 1989: however, the proportion of panel households (i.e. those sampled in at least two consecutive waves) has initially been fairly small, approaching 50% only in 1993 and 1995. Secondly, the structure of the questionnaire referring to the labour market varied considerably over time, and the 1993 and 1995 waves provide an acceptable degree of homogeneity in the available

[44] See Cannari and Gavosto [1994] for a full description of the subsection of the survey referring to the labour market.

information: as an example, the employer size, which we will see to have a considerable effect on the incidence of low-pay, is only available in the two selected waves. Finally, and probably most importantly, in 1993 a subsection on intergenerational mobility was introduced and, in particular, questions on the parents' education and occupation were asked to the spouse and the head of household: as will be clear later, such information plays a central role in the econometric analysis of earnings mobility and its absence from previous waves is the main reason which prevented the extension of the analysis to preceding transitions.

The characteristics of the data are reported in Table 4.1, where the first two columns refer to the sample composition in the 1993 and 1995 waves, while the third reports the same features, observed in 1993, for the panel sub-sample linking the two waves. The upper part of the Table illustrates the structure of the whole set of individual observations available under each partition; as we can see, the employees, both full and part-time and accounting for missing wage observations, amount at approximately one fourth of the sample, either in the two cross-sections and in the panel sub-sample. On the other hand, around 60% of the sample do not participate into the labour market.[45] By comparing the two cross-sections with the panel sub-sample, it can be observed how the proportions of students is slightly higher in the latter case, while the opposite is true for the retired, thus reflecting a higher propensity to stay within the household, and thus within sample, for students and inherently higher exit rates for pensioners.

The next panels in the Table go on to describe the sample structure for full-time employees with valid wage observations and aged between 18 and 65, which

[45] Given the well known importance of underground jobs in the Italian labour market, this is probably an overestimate. In the analysis which follows, I will consider only those employed on a regular basis and will not take into account individuals which, for example, report a labour income despite classifying themselves as retired.

will form the object of the econometric analysis. As can be seen, panel observations are now a smaller proportion relative to cross-sectional observations: the requirement for an observation to stay within the sample is now more demanding, which explains the fact. The differences in the sample composition between the cross-sections and the panel are not dramatic when age, experience and gender are taken into account, although in the first two cases the variable is slightly less disperse in the panel. A difference may instead be observed for what concerns the position in the household, the proportion of children in the panel sub-sample being some 4% lower than the two cross-sections, reflecting a higher propensity to leave the household in this group. Taking into account the other characteristics reported in the table, which basically consist of the wage determinants available in the SHIW data, we can see how, when compared with the two cross-sections, the panel sub-sample tends to be more educated, to hold non-manual jobs (teachers in particular), to be concentrated in the public administration[46] and to be employed in larger firms[47], all characteristics which indicate a stronger labour market attachment. This evidence suggests that panel attrition has an effect on the sample structure: the extent to which it affects estimates of the low-wage mobility model will be assessed in the next Chapter.

[46] The classification of sectoral affiliation in the SHIW questionnaire is jointly based on the type of product market and the public/private nature of the employer: this means that the coefficients on the public sector dummies in the next sections have to be interpreted not as public/private differentials, but as differentials between the public sector and the omitted category.
[47] Information on the employer's size only refers to private sector employees.

4.3. Definition and determinants of the low-pay status

This Section deals with the definition of the low pay threshold and with the quantification of the effect of observed workers characteristics on the probability of being low paid at a point in time.

A problem which is inherent to the analysis of low wage employment (and of poverty in general) is the definition of the threshold below which a worker may be considered a low wage earner. In particular, the problem is that of results robustness to the choice of the threshold. Various choices have been adopted in previous studies and, clearly, there are no *a priori* grounds to prefer one with respect to the others; to cope with this issue, here I follow the approach proposed by S&S and, instead of selecting a single threshold, I look at different thresholds in parallel.[48] In particular, I consider the first quintile and the third decile of the wage distribution of full time dependent workers aged between 18 and 65, which have both been used in previous studies (see Asplund *et al.* [1998] and Contini *et al.* [1998] respectively); both thresholds, being based on order statistics, guarantee robustness to outliers and avoid problems of updating over time.

A second issue is the definition of the wage variable. The wage information available in the SHIW data refers to the net annual wage, inclusive of overtime payments, and separately, to the monetary value of fringe benefits: for the purposes of the current analysis, I added them together to form the take-home net annual wage. This figure has then been normalised to account for heterogeneity in the amount of time effectively worked. Under this respect, the information available consists of the number of months effectively worked during the year and in the number of hours (inclusive of extra-time) averagely worked on a weekly basis; no information is available on the average number of weeks per month worked. This

[48] This strategy is also adopted by Jarvis and Jenkins [1997] in their analysis of low-income dynamics.

implies that to study hourly wages it is necessary to make some assumption on the number of weeks worked per month: here I follow Bardasi [1996] and assume that each individual worked 52/12 weeks each month. However, I also analyse monthly wages in parallel, so that any dramatic change in results between the two definitions can be checked.

Some features of the distribution of hourly and monthly (nominal) wages in the two years considered are reported in the upper panel of Table 4.2. As we can see, nominal wage growth has been fairly weak either at the mean and the median of the distribution for both wage measures, while wage dispersion has basically remained constant over the period. It can also be noted how the distribution of monthly wages tends to be more compressed, thus suggesting that heterogeneity in hours worked matters. The table also reports the low pay thresholds used in the analysis, and compares them with two thirds of the median wage, another threshold widely adopted in the literature; this last value tends to be lower than the first quintile. The lower panel of Table 4.2 deals with the proportions of workers which are defined low-paid under these three thresholds both in the cross-sectional sample and in the panel sub-sample. A first thing to note is that, in certain cases, the proportion of observations falling below or at a given percentile exceeds the level which one would expect from the percentile's definition, thus indicating the presence of clustering in the data. Secondly, we can observe how the lowest threshold (2/3 the median) is located around the fiftieth percentile for hourly wages and just above the first decile for monthly wages, again showing how this last variable is less dispersed. Finally, when the panel sub-sample is taken into account, the proportion of low-paid workers decreases under each threshold, a fact which is in line with the different structure of this group discussed above.[49]

A simple way of analysing the determinants of the low-pay status is to assess the effect of individual characteristics on the probability of being low-paid and to treat the problem by means of a discrete response model, namely a probit.[50] The use of discrete response models for the analysis of continuous variables clearly induces a loss of information. However, such models correspond to the idea that there exist a (possibly non-linear) change in the wage process below and above the low-pay threshold, i.e. that personal characteristics with a given impact at the middle or at the top of the distribution could produce rather different effects once the low-paid are taken into account. Moreover, while these models allow statements on the probability of low-pay without relying on any distributional assumption for wage levels or logarithms, some distributional assumption would be needed if one wanted to derive probability statements from a model for the continuous variable (as in Lillard and Willis [1978]). Finally, as already stressed in Chapter 1, the loss of information can be reduced by utilising discrete response models with multiple ordered outcomes, as I will do in Section 4.5.

Let us assume that, in a given year, wages depend on a set of individual and job characteristics:

$$g(w_i) = x_i'\delta + u_i \tag{4.1}$$

[49] In particular, this leads to small proportions for the lowest threshold, especially for monthly wages; this small cells problem was the reason which led to the exclusion of 2/3 the median from the econometric analysis. The same problem arises in OECD [1996].
[50] Probit regressions for the incidence of low-pay are estimated by Lucifora [1998] using the 1987 wave of the SHIW. The formalization used here is the one proposed by S&S.

where i indexes individuals, w is the nominal wage rate, x_i is a vector containing a constant and a set of wage determinants, δ is the vector of associated coefficients and $g(.)$ is a monotonic transformation such that u_i is standard normally distributed over i. Let λ be the low-pay threshold and d_i a dummy variable indicating the low-pay event: $d_i = \mathrm{I}(w_i \leq \lambda)$, where I(A) is a binary indicator which equals 1 when A is true and 0 otherwise.

Then, the probability that individual i will be low-paid is:

$$prob(d_i = 1) = prob(w_i \leq \lambda) = prob(g(w_i) \leq g(\lambda)) = \Phi(g(\lambda) - x_i'\delta) = \Phi(x_i'\beta) \quad (4.2)$$

where Φ is the standard normal cumulative distribution function (c.d.f.), the new constant term in β subsumes the difference between $g(\lambda)$ and the old constant in δ and the coefficients associated with the individual characteristics in β are the same as in δ, but with opposite sign.[51]

Such probit models for the low-pay probability have been estimated on the two SHIW cross-sections both for hourly and monthly wages; results are reported in Table 4.3.[52] Results are reported in terms of marginal effects, i.e. the change in predicted probabilities induced by a marginal change in the explanatory variable. For a continuous variable (say the j-th), these are evaluated as $\phi(\bar{x}'\hat{\beta})\hat{\beta}_j$ (where $\bar{x}$ is the vector of sample means of the explanatory variables and ϕ is the standard normal density function), while for a dummy variable (say the k-th) they're are computed as

[51] Given that this is a model for the probability of having a low wage, we should expect signs to revert with respect to a wage equation.
[52] The number of observations used in the estimation differs from the figures of Table 4.1 due to missing values in some of the explanatory variables. The same remark applies for the analysis of sections 4 and 5.

the change in predicted probabilities as the dummy changes from 0 to 1, all the other variables being evaluated at the sample mean, i.e. $\Phi(\hat{\beta}_k + \bar{x}_{-k}'\hat{\beta}_{-k}) - \Phi(\bar{x}_{-k}'\hat{\beta}_{-k})$, where the *-k* subscript denotes the corresponding vector deprived of the k-th element.

Looking first at each column of the table for hourly wages in isolation, it can be seen how the effect of personal characteristics tends to be in line with what one should expect from standard wage equations. Labour market experience (computed as age minus age at the beginning of the first job) has a non-linear effect on the probability of being low-paid, with the minimum located around 30 years. Educational qualifications have a negative impact on such a probability, with the effect of holding a BA degree which is roughly twice that of having an high school degree, both compared to those without an high school degree. Workers holding a non-manual job have a low-pay probability which is (depending upon the threshold) 10 to 26 percentage points lower when compared with blue collar workers; interestingly, the marginal effect for teachers is even higher than that for high level white collar workers, managers, university professors or magistrates[53], a fact which I will comment on later in the section. The effect of sectoral affiliation (with respect to manufacturing) is well determined for the public sector and agriculture, while the retail trade and services sectors display some effect depending upon the threshold or year considered; on the other hand, the employer size plays a clear role in reducing low-pay probabilities. Gender[54] and the region of residence have a significant effect; in particular, in the latter case it is the north-east which tends to have the lowest incidence of low-paid jobs. Finally, while both being married and

[53] Managers, professors and magistrates have been amalgamated with high level white collars because, since they tend not to fall below the threshold, a dummy for this group happens to be a "perfect classifier" and the corresponding parameter not identifiable.

[54] Rather than running a separate regression for each gender, I treat the effect with a dummy, in order to maintain homogeneity with the analysis of transition probabilities in the next section, where the pooling of female and male data has been necessary in order to preserve cells size.

head of the household significantly reduce the likelihood of low-pay, the presence of dependent (aged less than 14) children in the household has a less clear effect.

Taking now into account the estimates' stability over time, it can be noted how, apart from few exceptions, there are no dramatic differences. In particular, the size of the coefficients on the agriculture dummy drops considerably, while the effect for the retail trade group shows up only in the 1995 wave, which is also true (but only for the lower threshold) for the services sector. It is also interesting to observe how there is some evidence of geographical polarisation in low-pay probabilities over time, with the two northern marginal effects which tend to increase while the one for the centre falls.

Another interesting exercise is to control how estimated marginal effects change as the low-pay threshold is raised from the first quintile to the third decile. The general finding is that absolute values of significant effects tend to increase, while some effects which are non significant under the lower threshold become significant (this is the case for the services sector). This evidence is due to the fact that the bulk of observations which have personal characteristics with a given effect on the low-pay probability is located higher up in the wage distribution.[55]

The second part of Table 4.3 reports the results obtained for the distribution of monthly wages; differences with respect to hourly wages can then be ascribed to heterogeneity in hours supplied. Patterns emerged from the analysis of hourly wages are typically confirmed, but with some remarkable exception. First of all, the marginal effect for teachers is now the weaker (among occupational dummies) in absolute value, thus reversing the occupational ordering emerged from hourly wages. Secondly, a drop ranging from 3 to roughly 10 percentage points depending upon the

---

[55] In a separate experiment, I found that (for hourly wages in 1993) the effect of being a blue collar worker on the probability of having a wage below or at a given threshold grows monotonically until the median of the distribution and then falls.

threshold considered can be observed in the coefficients for the public sector. In both cases, heterogeneity in supply behaviour is determined by institutional factors. Finally, the female disadvantage in the probability of having a low wage is exacerbated in the monthly wage distribution, signalling that females tend to offer less hours than men, and the source of heterogeneity has more to do with behavioural factors.

4.4 The econometric analysis of low-wage transition probabilities

This Section takes advantage of the panel nature of the SHIW data to analyse the dynamics of the low-paid status at the individual level. The model proposed will enable detection of the workers' attributes relevant in determining low-pay persistence and, at the same time, of the forces driving falls into low-pay status from the upper part of the distribution, so that those personal characteristics which can guarantee the stability of higher hierarchical positions once reached can be identified. The extent of pure state dependence within aggregate persistence probabilities will also be analysed.

4.4.1 Aggregate transition probabilities

Before moving on to the econometric analysis of wage mobility, it may be instructive to look at the extent to which low-paid workers persist in their status at the aggregate level; such information is provided in Table 4.4, where raw transition probabilities from the 1993 to the 1995 status are reported both for hourly and monthly wages using the two low-pay definitions of the previous section; the first part of the table restricts the attention to the sample of employees in both years aged between 18 and 65 in 1993. The table points towards a substantial degree of low-pay persistence: 56% of those below the first quintile of hourly wages in 1993 are still low-paid in 1995, and such figure rises to nearly 71% when the threshold is defined in terms of the third decile. Similar figures, 61 and 64% respectively, arise for the monthly wage distribution. On the other hand, the probability of falling into low-pay from the top of the distribution is bounded below 10%.

These figures imply a considerable degree of (raw) state dependence in the conditional probability of being low-paid in 1995: if we use the difference $prob[L_{95}|L_{93}] - prob[L_{95}|H_{93}]$ (with L and H meaning low- and high-pay) as a measure of state dependence, we can see that it ranges from 50 to 60% depending upon the threshold and wage measure considered.

Although striking, such evidence may well imply different phenomena (Heckman [1981b]). On the one hand, it could be the result of workers heterogeneity, with the personal characteristics determining the low-pay status persisting over time; in this case, it is the difference in such characteristics between workers above and below the low-pay threshold which determines the observed state dependence. At the other extreme, raw figures may be generated by true state dependence, meaning that it is the experience of low-pay which modifies individual tastes or constraints and determines *per se* a higher persistence probability, holding fixed personal characteristics. As pointed out by S&S, true state dependence in low-pay persistence may arise from various models of the labour market. For example, if we think of low-paid jobs as "bad" jobs with no skill content, human capital models of wage determination can predict state dependence as a result of skill deterioration induced by the past experience of low-pay. The same prediction can arise in a signalling contest, where potential employers can use previous wages to make inference on the workers' quality and thus making low-wage offers to applicants who have formerly been low-paid. In addition, we could also think of a job search model where the experience of low-paid jobs induces workers to reduce their reservation wage, thus raising the probability of accepting low-wage offers in the future. The distinction between heterogeneity and true state dependence within aggregate persistence has relevant policy implications: while in the first case the probability of persisting in low-pay can be influenced by modifying workers' attributes (say via training programs), in the second more direct forms of low-wage protections are needed. Disentangling heterogeneity and true state dependence is thus an important issue in the analysis of low-pay transitions and the econometric analysis in this section will address this point.

Focusing only on those employed in both years could lead to ignore important aspects of the low-pay problem; for example, evidence of a cycle between low-pay and unemployment has been found for the UK (see Stewart [1999]). To shed light on the extent of the phenomenon in the SHIW data, Table 4.4 also considers transitions into other labour market states, namely self-employment, unemployment and retirement, for those aged 18 to 65 in 1993. In each of the four cases, the low-paid have a higher transition probability into both self-employment and unemployment when compared to the higher-paid, with raw state dependence being higher in the latter case. This suggests that low-wage jobs are characterised by a higher instability. On the other hand, a higher transition probability into retirement characterises the high-paid group, a likely effect of the life-cycle of earnings. Taking now into account the first column in each of the four matrices, we can also notice how in three out of four cases the unemployed are more likely to find a job below, rather than above, the low-pay threshold. This evidence is not enough to make statements about the existence of a cycle between unemployment and low-pay (which would require to observe at least two transitions), but is certainly not against such a hypothesis.

4.4.2 Model specification

The next step in this section is the construction of an econometric model of low-pay transition probabilities, i.e. the probability of being low-paid in 1995 conditional on the 1993 status; in particular, the object of the analysis will be the impact of personal characteristics, measured at the beginning of the transition[56], on individual transition probabilities. One central issue which arises in this context is that conditioning on the lagged state cannot be treated as exogenous: given that the wage process under investigation started prior to the sampling period its initial conditions are not observable by the researcher while, due to the presence of serial correlation in such a process, they will be embedded in wage levels at each time period, causing lagged wages to be endogenous with respect to current wages. This is the so-called *initial conditions problem* described in Heckman [1981a] and ignoring it can lead to biased estimates in the transition probability equation. The issue may also be thought of as a sample selection problem: if the propensity to be low-paid (or high-paid) in 1993 is not randomly distributed across the sample but depends on the unobservable initial conditions, estimating a transition equation selecting those who start from a low-pay (high-pay) state is endogenous to the transition probability.

This last remark suggests that some sort of correction for sample selection is needed; however, given the limited dependent nature of the transition equation, Heckman's correction techniques are not suitable in this context and the two probabilities (starting state and transition) have to be estimated jointly (O'Higgins [1994]).

[56] This qualification is aimed at avoiding endogeneity issues between changes in wages and changes in wage determinants.

To overcome the problem, here I extend the approach proposed by S&S[57] and treat it by means of a bivariate probit model with endogenous switching, i.e. the probit equivalent of usual endogenous switching models[58]. The model proposed by S&S assumes partial observability of the arrival wage distribution conditional on the origin wage distribution, i.e. for a given transition destination states are observable only for workers starting from low-pay. Such an hypothesis is not imposed by the lack of observations for those in a high-pay state at the beginning of the transition, but it is introduced for modelling purposes. I relax such an assumption and allow destination states to be observed also for the initially high-paid. This extension of the S&S's model thus implies a fuller exploitation of the information available. In this way, not only the effect of personal attributes on low-pay persistence can be estimated, but also the impact of these same factors on the probability of falling into low-pay from the top of the distribution can be assessed within the same model.[59] This in turn allows assessment of the extent with which the effect of observable attributes on low-pay transitions varies with the starting state.

Let us specify the selection equation for the initial state along the lines adopted in Section 4.3 to model the low-pay probability at a point in time:

$$\begin{aligned} g(w_{i93}) &= x_i'\delta + u_i \\ d_i &= \mathrm{I}(w_{i93} \le \lambda_{93}) \end{aligned} \qquad (4.3)$$

where the specification of the x-vector differs from Table 4.3, as will be clear later.

[57] The S&S's model corresponds to model no. 3 in Meng and Schmidt [1981] catalogue of bivariate probit models; an application of this type of model is given in Section 4.5.

[58] Endogenous switching equations models for continuous variables are set out in Lee [1978].

[59] If one wanted to analyse the two types of effects with the S&S's model, estimation of two models would be needed.

Next, suppose that the effect of exogenous variables on the arrival state depends upon the initial state in the following way:

$$\begin{array}{lll} h_1(w_{i95}) = z_i'\eta_1 + \varepsilon_{1i} & if & d_{i93} = 1 \\ h_2(w_{i95}) = z_i'\eta_2 + \varepsilon_{2i} & if & d_{i93} = 0 \end{array}, \qquad (4.4)$$

where $h_j(.)$ is a monotonic transformation such that ε_{ji} is standard normally distributed over individuals and z is a subvector of x.[60] Let d_{i95} be a dummy variable indicating the low-pay event in the arrival wage distribution and assume that u and the ε's are jointly distributed as a tri-variate normal:

$$\begin{pmatrix} u_i \\ \varepsilon_{1i} \\ \varepsilon_{2i} \end{pmatrix} \sim N_3\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & & \\ \rho_1 & 1 & \\ \rho_2 & \rho_3 & 1 \end{pmatrix}\right]^{61}.$$

Given the assumptions on the error distributions, it follows that:

$$\begin{array}{l} prob(d_{i95} = 1, d_{i93} = 1) = \Phi_2(z_i'\gamma_1, x_i'\beta; \rho_1) \\ prob(d_{i95} = 1, d_{i93} = 0) = \Phi_2(z_i'\gamma_2, -x_i'\beta; -\rho_2) \end{array}, \qquad (4.5)$$

where Φ_2 is the bivariate normal c.d.f., β derives from δ in the same fashion of Section 4.3 and analogously for the γ's and η's; thus the elements of γ_1 model the effect of individual characteristics on low-pay persistence, while γ_2 captures the effect of the same characteristics on the probability of falling from the upper part of the distribution into low-pay. Note that although these expressions refer to the joint

[60] The S&S's model assumes that w_{i95} is not observable if d_{i93}=0.

[61] Note that ρ_3 is not identifiable since it would require observations belonging contemporaneously to both regimes.

probability, estimation of the γ_j's is based on sub-samples defined according to the starting state and is, in this sense, conditional. Note also that, given the model's structure, only the evaluation of the bivariate normal cdf is required. To derive the correct (i.e. summing to one over the sample of the initially low or high-paid) expression for the conditional probability we need to normalise on the probability of the initial state:

$$\begin{aligned} prob(d_{i95} = 1 | d_{i93} = 1) &= \frac{\Phi_2(z_i'\gamma_1, x_i'\beta; \rho_1)}{\Phi(x_i'\beta)} \\ prob(d_{i95} = 1 | d_{i93} = 0) &= \frac{\Phi_2(z_i'\gamma_2, -x_i'\beta; -\rho_2)}{\Phi(-x_i'\beta)}, \end{aligned} \tag{4.6}$$

which makes clear how the parameters for such transition probabilities can be consistently estimated with a univariate probit on sub-samples defined according to the starting state only if ρ_j=0, i.e. only if the starting state is exogenous.

The log-likelihood function of the model may be written as:

$$\begin{aligned} \log L = \sum_i \{ & d_{i93} d_{i95} \log[\Phi_2(z_i'\gamma_1, x_i'\beta; \rho_1)] + d_{i93}(1 - d_{i95}) \log[\Phi_2(-z_i'\gamma_1, x_i'\beta; -\rho_1)] + \\ & (1 - d_{i93}) d_{i95} \log[\Phi_2(z_i'\gamma_2, -x_i'\beta; -\rho_2)] + \\ & (1 - d_{i93})(1 - d_{i95}) \log[\Phi_2(-z_i'\gamma_2, -x_i'\beta; \rho_2)] \}. \end{aligned} \tag{4.7}$$

Identification of the transition process in (4.4) requires restrictions in the form of variables which enter the x-vector but not the z-vector; in the present case we need variables which influence the wage level but, given this, have no direct effect on the wage change. Here I adopt S&S's identification strategy and use a set of indicators of the worker' s parental background in terms of her parents education and occupation. As stated in Section 4.2, since 1993 the SHIW questionnaire contains a part on intergenerational mobility, where the spouse and head of household are asked to report, among others, their parents' education and occupation. For those workers who are "children" in the interviewed household, the necessary information has directly been recovered from the household questionnaire. Going back to Table 4.2, this means that for 1.58% of the estimation sample (i.e. those who are "other relative or non-relative" in the interviewed household) such parental background variables are not available. In order to preserve the sample size, I treated these cases and the ones where the parental background information was "genuinely" missing with dummies for missing information.[62]

Besides the parental background indicators, another variable which only enters the selection equation is the square of labour market experience, given the interpretation of wage change equation which can be attributed to (4.4), i.e. states in 1995 conditional on states in 1993. This implies that the equation for the transition probability is over-identified and that the validity of the parental background variables as instruments can be tested: such tests are presented along with the estimation results.

4.4.3 Results

Before considering the whole set of results from the switching probit analysis, Table 4.5 compares estimated ML coefficients under the two competing assumptions (i.e. endogeneity *versus* exogeneity) on the conditioning starting state, focusing, for expositional compactness, on the low-pay threshold defined as the bottom quintile of the hourly wage distribution.[63] The table gives a flavour of the kind of bias induced by

[62] These are, typically, negligible proportions of the sample, reaching at most 4%; only in the case of the mother's occupation the figure rises to 14%.
[63] Results similar to the ones reported were obtained for the other low-pay and wage definitions. The exogenous starting state estimates are probit models for the 1995 low-pay event estimated on sub-

assuming exogenous initial conditions. First of all it can be noticed that the null hypothesis of exogenous starting state is rejected for both starting states (i.e. low-pay and high-pay), the two correlation coefficients being statistically significant at conventional levels. Taking estimated coefficients into account, it can be observed that the exogeneity hypothesis leads to overestimate both their size and significance. This is true especially in the case of labour market experience, whose effect on the conditional probability of being low-paid vanishes once allowance is made for endogeneity. For the remaining explanatory variables such overestimation is, although less pronounced, also evident; on the whole, results from Table 4.5 confirm similar comparisons reported by S&S and warn against the dangers of assuming exogeneity of initial conditions.

Results from the switching bivariate probit model are given in Table 4.6 for each low-pay threshold and wage definition, both in terms of ML coefficients and associated marginal effects on the conditional probability.[64] By considering correlation coefficients first, it can be observed how, in each case, they are statistically significative at usual confidence levels, thus clearly rejecting the hypothesis of initial conditions' exogeneity. Such parameters are negative; given that

---

samples defined according to the 1993 position in the wage distribution, i.e. above or below the low-pay threshold.

[64] For each explanatory variable, the marginal effect is given first, followed by the ML estimated coefficient and the asymptotic t-ratios. The computation of marginal effects from the bivariate probit estimates requires some additional caution, given that a change in a variable in z implies also a change in the corresponding element of x and thus in the denominator of the conditional probability. What we would require is instead a change in the conditional probability holding the past fixed (see S&S). With this aim, and focusing for the exposition's sake on the probability of low-pay persistence, let us define the average predicted probability of initial low-pay as $\hat{\Phi} = \sum_i \Phi(x_i'\hat{\beta}) / N$ (N is the sample size) and its argument as $\widehat{x\beta} = \Phi^{-1}(\hat{\Phi})$; the marginal effect for the k-th dummy variable is then computed as $\frac{\Phi_2(\hat{\gamma}_{1k} + \bar{z}_{1-k}'\hat{\gamma}_{1-k}, \widehat{x\beta}; \rho_1)}{\hat{\Phi}} - \frac{\Phi_2(\bar{z}_{1-k}'\hat{\gamma}_{1-k}, \widehat{x\beta}; \rho_1)}{\hat{\Phi}}$, z1 indicating that the average is taken over the relevant sample, the initially low-paid in this case. For labour market experience the effect has been computed as that of a discrete change from 20 to 30 years of experience.

they measure the correlation between the probability of having a small wage change and the probability of having a low initial wage, the negative sign is analogous to a negative coefficient estimated in the regression of wage changes on wage levels, i.e. Galtonian regression towards the mean. Note also that, given the structure of the model and, in particular, the uniqueness of the selection equation which models the probability of having an initial low-wage, this is true also for the initially high-paid. Another fact to note is that the identifying restrictions on the parental background variables are supported by the data at usual confidence levels.

Taking the effect of observable characteristics into account, it can be noticed how labour market experience has basically no effect in reducing the conditional probability of having a low-wage. Educational qualifications, on the other hand, have an effect in such direction which tends to be stronger for those starting the transition below the low-pay threshold; the same is true for the female dummy, but with opposite sign. Non-manual jobs and jobs in large firms are instead characteristics which tend to prevent workers from falling into low-pay, while the effect on low-pay persistence is less robust; similar considerations, but only for the hourly wage distribution, apply for the public sector dummy. The agricultural sector dummy seems to favour drops into low-pay for the distribution of monthly wages, thus denoting a certain wage instability for these jobs. On the other hand, holding a job in the service sector positively affects low-pay persistence[65], while no effect is detected on drops from the high-pay area. Such a result could arise from those workers who, say in a bank or an insurance company, are on a low-level job career (actually involving manual tasks such as delivering) but do not classify themselves as blue collar. An alternative explanation could be that this service category is broad enough to include

[65] Similar results on Italian data are reported by Contini et al. [1998].

cases which markedly differ from the conventional perception of service sector. Finally, the geographical dummy is significative in reducing low-pay persistence, while no effect can be detected for those initially high-paid.

As we saw earlier in this section, one important issue in the dynamic analysis of low-pay is the distinction between true state dependence and heterogeneity within raw persistence probabilities. Estimation results from Table 4.6 enable such a decomposition, which is reported in the last three rows of the table. The row labelled "Estimated state dependence" reports the difference in the conditional probability of being low-pay computable from the estimated model, giving a measure of overall state dependence which is, apart from small differences due to observations with missing values in the explanatory variables excluded from the regression analysis, the same as the aggregate state dependence effect of Table 4.4:

$$ESD = \frac{\sum_{i:di93=1} \dfrac{\Phi_2(z_i \hat{\gamma}_1, x_i \hat{\beta}; \hat{\rho}_1)}{\Phi(x_i \hat{\beta})}}{\sum_i d_{i93}} - \frac{\sum_{i:di93=0} \dfrac{\Phi_2(z_i \hat{\gamma}_2, -x_i \hat{\beta}; -\hat{\rho}_2)}{\Phi(-x_i \hat{\beta})}}{\sum_i (1-d_{i93})}. \tag{4.8}$$

The measure of true state dependence has been obtained by computing this same quantity but holding fixed the sample over which it is averaged, i.e. abstracting from heterogeneity in observable explanatory variables between workers below and above the low-pay threshold in the origin wage distribution. This procedure yields two measures of true state dependence, corresponding to the two sub-samples over which the average is taken, which are reported in the two bottom lines. Such measures are equivalent to "price" effects in a classical Oaxaca decomposition of wage differentials; in terms of true state dependence, the "price" effect captures the extent to which workers with the same observable characteristics are evaluated differently according to their past wage, i.e. the parameters of their environment are changed by the past low-pay experience per se. First of all it can be observed how true state dependence constitutes a considerable share of aggregate state dependence, ranging from 40 to 70%, thus indicating that, to a meaningful extent, the sole experience of low-paid jobs increases the likelihood of experiencing low-wages in the future, as could result from the presence of low-pay stigma, human capital depreciation or changes in search behaviour, all factors which, as stressed above, could be the way by which the occurrence of low-pay induces alterations of the economic environment. Secondly, true state dependence is higher when the parameters estimates are applied to the sample of the initially low-paid, probably as an effect of unobserved heterogeneity between workers above and below the low-pay threshold, not captured by the controls adopted in the model.

4.5 Accounting for the width of transitions

As is well recognised by the statistical literature on mobility (see, for example, Boudon [1972]), an important feature of the mobility process is given by the magnitude of the "jumps" made by those workers abandoning the origin wage class: not only the fact of changing wage rank is important, but also the width of such transitions matters in assessing the degree of distributional mobility. This is the reason why the descriptive literature on mobility indices reviewed in Chapter 1 developed measures such as the *expected absolute jump*. In terms of the econometric modelling of transition probabilities, accounting for their width can give some indication on the loss of information induced by the dichotomic treatment of the wage variable underlying the switching bivariate probit above. In other words, the model of Section 4.4 considers only one alternative to the low-pay status in the destination wage distribution, and some of the effects significant in affecting low-pay persistence may well result from small wages "pushes", just sufficient to bring individuals above the low-pay threshold. If one views these small pushes as potential sources of measurement errors, then the model utilised in this Section allows a more robust assessment of the forces governing the transition process.

To get a feeling on the extent of upward movements from the low-pay status, I report below the aggregate transition probabilities from the bottom three deciles of the distribution.

Transition probabilities from the bottom three deciles of the wage distribution (N=2160)

hourly	1	2	3	4	5	6	7	8	9	10	% 1993
1	44.25	26.44	15.52	3.45	4.60	1.72	2.30	1.15	0.57	0.00	8.06
2	17.58	23.03	26.06	9.70	14.55	3.03	2.42	1.21	2.42	0.00	7.64
3	6.15	18.46	35.90	11.79	13.33	6.15	4.10	2.05	1.03	1.03	9.03
monthly											
1	44.23	28.21	10.26	5.77	1.28	4.49	0.64	4.49	0.00	0.64	7.22
2	18.48	34.24	13.59	16.30	7.07	7.61	1.09	0.54	1.09	0.00	8.52
3	9.19	23.78	15.14	22.70	11.89	8.65	3.24	3.78	0.54	1.08	8.56

As we can see there's considerable variation in the destination states of those who cross the low-pay threshold, and while the bulk of transitions reach just the decile just adjacent the low-pay area, there are some cases (in particular starting from the third decile) in which the median of the distribution is crossed.

A way to investigate the impact of transition width on the parameters of interest is to allow for more than two outcomes in the transition equation. With this aim, I propose a second extension of the S&S's model. In particular, I use their model with partial observability (1995 outcomes observable only for the 1993 low-paid) but model the transition equation with an ordered probit, rather than with a binary probit as in their case.[66] Let us assume that selection into the starting state is still governed by (4.3), while the position in the destination wage distribution can only be observed for the initially low-paid (i.e., only the first part of (4.4) applies) and is represented by the following discrete ordered indicator:

$$d_{i95} = \begin{cases} 1 & \text{if} & & w_{i95} \leq & \lambda_{95} \\ 0 & \text{if} & \lambda_{95} & < w_{i95} \leq & \lambda_{95} + \mu_0 \\ -1 & \text{if} & \lambda_{95} + \mu_0 & < w_{i95} \leq & \lambda_{95} + \mu_1 \\ -2 & \text{if} & \lambda_{95} + \mu_1 & < w_{i95} \leq & \lambda_{95} + \mu_2 \\ -3 & & & \text{otherwise} & \end{cases} \tag{4.9}$$

where the μ's are the first three deciles above the low-pay threshold, while the assumptions on the joint distribution of u_i and ε_{1i} are unaltered[67]. The resulting log-likelihood is:

$$\begin{aligned} \log L = \textstyle\sum_i \{ & \mathrm{I}(d_{i95} = 1) d_{i93} \log[\Phi_2(z_i\gamma_1, x_i\beta; \rho_1)] + \\ & \mathrm{I}(d_{i95} = 0) d_{i93} \log[\Phi_2(\nu_0 + z_i\gamma_1, x_i\beta; \rho_1) - \Phi_2(z_i\gamma_1, x_i\beta; \rho_1)] + \\ & \mathrm{I}(d_{i95} = -1) d_{i93} \log[\Phi_2(\nu_1 + z_i\gamma_1, x_i\beta; \rho_1) - \Phi_2(\nu_0 + z_i\gamma_1, x_i\beta; \rho_1)] + \\ & \mathrm{I}(d_{i95} = -2) d_{i93} \log[\Phi_2(\nu_2 + z_i\gamma_1, x_i\beta; \rho_1) - \Phi_2(\nu_1 + z_i\gamma_1, x_i\beta; \rho_1)] + \\ & \mathrm{I}(d_{i95} = -3) d_{i93} \log[\Phi(x_i\beta) - \Phi_2(\nu_2 + z_i\gamma_1, x_i\beta; \rho_1)] + (1 - d_{i93}) \log[\Phi(-x_i\beta)] \}, \end{aligned} \tag{4.10}$$

where the ν_j's ($=h_1(\lambda_{95}+\mu_j)$) are parameters to be estimated.

[66] Guillotin and Hamouche [1998] model the number of jumps by means of count data models in a framework with exogenous initial conditions.

[67] The specification in (4.9) is aimed at maintaining the comparability of $\gamma 1$ with the analogous vector estimated from the analysis in Section 4.4.

Results from the estimation of this ordered probit with selectivity are reported in Table 4.7 and are compared with those from a switching bivariate probit with partial observability, i.e. where the polychotomous indicator in (4.9) is replaced by a binary indicator (1 for w_{i95} below the low-pay threshold and 0 otherwise). A first thing to note is that in each of the cases considered, the null of exogenous initial conditions is rejected at conventional levels, while the validity of the parental background indictors as instruments for the starting state is supported by the data. By comparing the correlation coefficient across the ordered and binary probit models, it can be observed that it is always lower (bigger in absolute value) in the first case. If we recall that a negative value of this parameter reflects the fact that small wage gains are negatively associated with low initial wages, its behaviour across models suggests that in the polychotomous framework low-pay persistence is a relatively worse outcome than in the binary case. On the other hand, for statistically significant coefficients and associated marginal effects[68] the general finding (a remarkable exception is the dummy for the service sector) is that they decrease in absolute value as we move from the binary to the polychotomous specification of the position in the 1995 wage distribution, meaning that part of such effects was due to small wage "pushes".

4.6 Summary and conclusions

This Chapter has utilised panel data from the 1993 and 1995 waves of the Bank of Italy's SHIW to analyse the determinants of low-wage mobility.

[68] As for the preceding analysis, such effects refer to variations in the probability of being low-paid in 1995 conditional on low-pay in 1993. Their computation therefore coincides with the one reported in note 64 and, in particular, the conditioning probability is still given by a binary probit for low-pay in 1993. The relevant difference is that now the estimated γ_1 reflects the existence of more than one alternative to the low-pay status in 1995.

Defining the low-paid alternatively as those below the bottom quintile or the third decile of the wage distribution, both in hourly and monthly terms, the usual set of wage determinants (human capital, demand side and demographic variables) has been found to have a significant effect on the probability of being low-paid at a point in time.

The analysis has next turned to low-pay dynamics at the individual level. The econometric analysis of low-wage mobility has been based on a bivariate probit model with endogenous switching, which extends the approach previously proposed by Stewart and Swaffield [1999] for the assessment of the initial conditions problem, i.e. the potential endogeneity of the initial low-pay status.

Results show how the hypothesis of exogenous initial conditions can always be rejected, the correlation coefficient between the unobservables in the starting state and transition equations being significantly different from zero. By comparing these results with those from models where the initial status is taken as exogenous, the paper has shown how in this last case the effects of mobility determinants are systematically overstated both in size and significance: this is especially true for labour market experience. Among the other variables controlled for, education, gender, sectoral affiliation to the service sector and geographical location have been found to affect low-pay persistence, while non-manual occupations and jobs in large firms are effective in avoiding falls into low-pay once higher wage positions have been reached, while their effects on low-pay persistence appear to be less robust. This last remark applies also to affiliation to the public sector, but only for the hourly wage distribution.

Estimates from the endogenous switching bivariate probit have been utilised to assess the extent of *true* state dependence within raw transition probabilities: it has been shown that true state dependence affects wage profiles to a meaningful extent, between 40 and 70% of raw state dependence: thus, a considerable share of low-pay persistence appears to arise from the experience of low-pay *per se*, rather than from heterogeneity in persistent observable workers attributes above and below the low-pay threshold.

Some attempt has also been made to understand the consequences of the binary treatment of the wage variable underlying the endogenous switching model. Using an ordered probit model with endogenous sample selection, it has been shown how, typically, significant effects tend to drop in size, suggesting that their effectiveness is to some extent confined to the quantiles just adjacent to the low-pay threshold.

These results show that while factors which are traditionally known as wage determinants have a limited effect on the conditional probability of abandoning the low-pay status, the past experience of low-pay has, *per se*, a considerable impact on future low-pay probabilities, both circumstances which raise concern about the welfare of workers at the bottom of the wage distribution. However, data limitations, in particular the fact that a single transition has been analysed, suggest caution in drawing conclusions and prompt future research on this issue.

4. Tables

Table 4.1: Sample description

	1993		1995		panel (1993)	
	n.obs/mean	%/s.d.	n.obs/mean	%/s.d.	n.obs/mean	%/s.d.
employed	5768	24.02	5598	23.4	2634	24.49
employed missing wage/part time	468	1.95	578	2.42	206	1.92
self-employed	1302	5.42	1492	6.24	564	5.24
entrepreneurs	585	2.44	557	2.33	256	2.38
seek first job	1215	5.06	1029	4.3	487	4.53
unemployed	511	2.13	690	2.88	234	2.18
retired	5401	22.49	5448	22.77	2193	20.39
student	4528	18.86	4400	18.39	2274	21.14
housewife	1247	5.19	1284	5.37	570	5.3
other	2988	12.44	2848	11.9	1337	12.43
Total	24013	100	23924	100	10755	100
Employed 18<=age<=65	5708		5541		2160	
age	39.0555	10.7613	38.9821	10.7944	39.4032	9.99305
experience/10	1.92771	1.14947	1.96128	1.16041	1.90736	1.06188
male	3677	64.42	3510	63.35	1391	64.4
female	2031	35.58	2031	36.65	769	35.6
head of family	2998	52.52	2794	50.42	1191	55.14
spouse/cohabitee	1312	22.99	1328	23.97	536	24.81
child	1265	22.16	1310	23.64	399	18.47
other relative-non relative	133	2.33	109	1.97	34	1.58
no school	81	1.42	55	0.99	16	0.74
elem. school (5 yrs)	824	14.44	679	12.25	260	12.04
junior high (8 yrs)	2035	35.65	2137	38.57	707	32.73
high school (13 yrs)	2123	37.19	2019	36.44	879	40.69
ba/bs (17+ yrs)	645	11.3	651	11.75	298	13.8
blue collar	2515	44.06	2525	45.57	854	39.54
white collar low level	2168	37.98	1814	32.74	832	38.52
teacher	598	10.48	647	11.68	293	13.56
white collar high level	289	5.06	416	7.51	124	5.74
manag,professor,magistrate	138	2.42	139	2.51	57	2.64
agriculture	164	2.87	133	2.4	42	1.94
other manufacturing	1597	27.99	1700	30.68	578	26.76
construction	333	5.84	300	5.41	105	4.86
retail trade	524	9.18	529	9.55	168	7.78
transport & communication	161	2.82	174	3.14	50	2.31
bank insurance	191	3.35	208	3.75	76	3.52
real estate	166	2.91	138	2.49	67	3.1
domestic & other services	191	3.35	188	3.39	59	2.73
public administration	2379	41.69	2171	39.18	1015	46.99
size<=4	470	14.12	484	14.36	145	12.45
5<=size<=19	930	27.94	950	28.19	272	23.35
20<=size<=49	499	14.99	476	14.12	163	13.99
50<=size<=99	324	9.73	280	8.31	107	9.18
100<=size<=499	474	14.24	499	14.81	179	15.36
size>=500	632	18.98	681	20.21	299	25.67
northwest	1389	24.33	1400	25.27	503	23.29
northeast	1200	21.02	1273	22.97	494	22.87
centre	1313	23	1162	20.97	426	19.72
south	1322	23.16	1235	22.29	533	24.68
islands	484	8.48	471	8.5	204	9.44

Table 4.2: Descriptive statistics of the wage distribution (upper panel) and incidence of low-pay for different thresholds (lower panel)

	hourly wages		monthly wages	
	1993	1995	1993	1995
Descriptive statistics (thousands of lire)				
mean	12.36	12.86	1958.30	2061.47
median	10.82	11.54	1833.33	1916.67
sd logs	0.47	0.45	0.40	0.39
log(90/10)	1.06	1.05	0.87	0.88
2/3 median	7.22	7.69	1222.22	1277.78
first quintile	8.05	8.55	1375.00	1500.00
third decile	8.97	9.62	1500.00	1625.00
Low-pay incidence				
2/3 median	14.02	15.03	11.3	11.19
2/3 median (panel)	10.28	9.58	7.82	6.3
bottom quintile	20.04	20.43	20.08	24.36
bottom quintile (panel)	15.69	14.07	15.74	16.57
third decile	30.2	34	30.17	30.63
third decile (panel)	24.72	24.77	24.31	22.55

Table 4.3: Probit marginal effects for the probability of being low-paid: hourly wages

Threshold	Bottom Quintile				Third Decile			
	1993		1995		1993		1995	
experience/10	-0.133	(9.59)	-0.108	(7.49)	-0.192	(9.00)	-0.215	(9.28)
experience^2/100	0.024	(7.66)	0.018	(5.92)	0.034	(7.39)	0.036	(7.30)
high school degree	-0.044	(3.64)	-0.045	(3.67)	-0.089	(5.00)	-0.091	(4.68)
ba degree +	-0.085	(4.41)	-0.083	(3.99)	-0.176	(6.17)	-0.187	(5.87)
white collar low-level	-0.097	(8.23)	-0.093	(7.74)	-0.165	(9.65)	-0.188	(10.14)
teachers	-0.103	(5.94)	-0.099	(4.96)	-0.213	(8.17)	-0.260	(9.07)
white collar high level, managers, uni. prof., magistrate	-0.091	(4.63)	-0.101	(5.72)	-0.176	(6.07)	-0.219	(8.31)
public sector	-0.189	(14.20)	-0.202	(14.14)	-0.317	(16.25)	-0.326	(15.04)
agriculture,forests	0.147	(5.08)	0.059	(2.10)	0.140	(3.41)	0.040	(0.91)
construction	-0.016	(1.04)	0.019	(1.04)	-0.016	(0.64)	0.011	(0.39)
retail trade, household & other services	0.013	(0.98)	0.034	(2.49)	-0.002	(0.11)	0.059	(2.57)
transport & comm.	-0.030	(1.23)	-0.024	(0.97)	-0.030	(0.82)	-0.056	(1.42)
bank, insurance, real estate	0.004	(0.20)	-0.034	(1.70)	-0.061	(2.17)	-0.083	(2.63)
20<=firm size<=99	-0.077	(8.55)	-0.059	(5.69)	-0.132	(8.40)	-0.100	(5.18)
100<=firm size<=499	-0.090	(8.46)	-0.103	(9.56)	-0.174	(9.78)	-0.178	(8.73)
size>=500	-0.118	(11.38)	-0.119	(10.86)	-0.213	(12.69)	-0.245	(13.06)
female	0.096	(8.65)	0.087	(7.58)	0.140	(8.55)	0.137	(7.71)
north-west	-0.068	(6.53)	-0.096	(9.00)	-0.098	(5.88)	-0.100	(5.37)
north-east	-0.080	(7.79)	-0.093	(8.91)	-0.103	(6.06)	-0.133	(7.17)
centre	-0.055	(5.27)	-0.048	(4.22)	-0.064	(3.79)	-0.049	(2.50)
married	-0.074	(5.55)	-0.048	(3.61)	-0.129	(6.79)	-0.077	(3.79)
head of household	-0.050	(4.41)	-0.038	(3.20)	-0.078	(4.69)	-0.076	(4.19)
dependent children	0.015	(1.25)	-0.007	(0.55)	-0.010	(0.60)	-0.053	(2.90)
Number of obs		5673		5522		5673		5522
chi2(23)		2061.13		1885.99		2494.03		2557.12
Prob > chi2		0		0		0		0
Pseudo R2		0.3642		0.3378		0.3594		0.3613

Table 4.3 (continued): Probit marginal effects for the probability of being low-paid: monthly wages

Threshold	Bottom Quintile				Third Decile			
	1993		1995		1993		1995	
experience/10	-0.150	(10.55)	-0.192	(10.99)	-0.219	(10.34)	-0.276	(13.11)
experience^2/100	0.027	(8.41)	0.033	(8.76)	0.038	(8.24)	0.048	(10.67)
high school degree	-0.066	(5.26)	-0.072	(4.74)	-0.112	(6.26)	-0.114	(6.30)
ba degree +	-0.099	(5.69)	-0.115	(5.07)	-0.190	(7.41)	-0.176	(6.71)
white collar low-level	-0.099	(7.95)	-0.119	(7.92)	-0.204	(11.85)	-0.165	(9.33)
teachers	-0.072	(3.80)	-0.096	(4.06)	-0.171	(6.66)	-0.134	(4.79)
white collar high level, managers, univ. professor, magistrate	-0.095	(4.43)	-0.154	(7.03)	-0.210	(7.26)	-0.213	(8.26)
public sector	-0.156	(11.20)	-0.174	(9.93)	-0.221	(10.85)	-0.188	(8.85)
agriculture,forests	0.147	(4.85)	0.049	(1.43)	0.142	(3.41)	0.049	(1.20)
construction	-0.008	(0.44)	0.012	(0.52)	-0.018	(0.71)	0.033	(1.15)
retail trade, household & other services	-0.002	(0.13)	0.014	(0.81)	-0.010	(0.46)	0.010	(0.50)
transport & comm.	-0.049	(1.91)	-0.045	(1.36)	-0.037	(0.96)	-0.086	(2.21)
bank, insurance, real estate	0.009	(0.41)	-0.053	(2.07)	-0.013	(0.41)	-0.059	(1.83)
20<=firm size<=99	-0.077	(7.64)	-0.067	(4.69)	-0.117	(6.97)	-0.084	(4.60)
100<=firm size<=499	-0.097	(8.27)	-0.136	(9.09)	-0.167	(8.67)	-0.164	(8.44)
size>=500	-0.124	(10.47)	-0.159	(10.33)	-0.211	(11.41)	-0.214	(11.18)
female	0.137	(11.57)	0.153	(10.87)	0.200	(11.99)	0.194	(11.70)
north-west	-0.063	(5.77)	-0.100	(7.40)	-0.106	(6.45)	-0.084	(4.95)
north-east	-0.074	(6.77)	-0.108	(7.96)	-0.121	(7.27)	-0.083	(4.84)
centre	-0.056	(5.11)	-0.054	(3.75)	-0.077	(4.64)	-0.025	(1.41)
married	-0.083	(6.08)	-0.088	(5.49)	-0.120	(6.41)	-0.094	(5.06)
head of household	-0.054	(4.63)	-0.058	(4.11)	-0.075	(4.53)	-0.060	(3.62)
dependent children	0.005	(0.36)	-0.007	(0.45)	-0.026	(1.51)	-0.028	(1.62)
Number of obs		5673		5522		5673		5522
chi2(23)		1929.88		1847.26		2288.77		2074.12
Prob > chi2		0		0		0		0
Pseudo R2		0.3405		0.3015		0.3301		0.305

Notes: asymptotic t-ratios in parentheses, reference categories are no or elementary education, blue collar, manufacturing, firm size<=19, male, residence in the south or isles, not married, not head of household, no dependent children in the household

Table 4.4: Aggregate transition probabilities between labour market states (L=low-pay, H=high pay, SE=self employment, UN=unemployment, RET=retired)

Hourly wages

threshold=bottom quintile

N=2160	L 95	H 95	% 93
L 93	56.05	43.95	15.69
H 93	6.26	93.74	84.31
% 95	14.07	85.93	

threshold=third decile

N=2160	L 95	H 95	% 93
L 93	70.79	29.21	24.72
H 93	9.66	90.34	75.28
% 95	24.77	75.23	

Monthly wages

threshold=bottom quintile

N=2160	L 95	H 95	% 93
L 93	61.76	38.24	15.74
H 93	8.13	91.87	84.26
% 95	16.57	83.43	

threshold=third decile

N=2160	L 95	H 95	% 93
L 93	64.76	35.24	24.31
H 93	8.99	91.01	75.69
% 95	22.55	77.45	

Hourly wages, L=bottom quintile

N=4096	L 95	H 95	SE 95	UN 95	RET 95	%93
L 93	46.91	36.79	3.21	10.37	2.72	9.95
H 93	5.59	83.76	1.08	2.11	7.46	50.09
SE 93	0.81	1.41	88.48	3.43	5.86	12.17
UN 93	13.02	13.02	9.38	59.90	4.69	4.72
RET 93	0.00	0.21	0.85	0.32	98.62	23.08
% 95	8.18	46.45	12.26	5.41	27.70	

Hourly wages, L=third decile

N=4096	L 95	H 95	SE 95	UN 95	RET 95	%93
L 93	59.72	24.64	2.69	9.32	3.63	15.56
H 93	8.67	81.16	0.99	1.44	7.73	44.48
SE 93	1.01	1.21	88.48	3.43	5.86	12.17
UN 93	18.75	7.29	9.38	59.90	4.69	4.72
RET 93	0.00	0.21	0.85	0.32	98.62	23.08
% 95	14.16	40.48	12.26	5.41	27.70	

Monthly wages, L=bottom quintile

N=4096	L 95	H 95	SE 95	UN 95	RET 95	%93
L 93	51.47	31.86	3.68	10.29	2.70	10.03
H 93	7.27	82.16	0.98	2.11	7.47	50.01
SE 93	0.61	1.62	88.48	3.43	5.86	12.17
UN 93	17.71	8.33	9.38	59.90	4.69	4.72
RET 93	0.00	0.21	0.85	0.32	98.62	23.08
% 95	9.71	44.93	12.26	5.41	27.70	

Monthly wages, L=third decile

N=4096	L 95	H 95	SE 95	UN 95	RET 95	%93
L 93	54.14	29.46	2.71	9.71	3.98	15.43
H 93	8.10	81.98	0.99	1.32	7.60	44.61
SE 93	0.81	1.41	88.48	3.43	5.86	12.17
UN 93	20.83	5.21	9.38	59.90	4.69	4.72
RET 93	0.00	0.21	0.85	0.32	98.62	23.08
% 95	13.05	41.58	12.26	5.41	27.70	

Table 4.5. Comparison of ML estimates of conditional low-pay probabilities equations under competing assumptions on initial conditions. Bottom quintile of the hourly wage distribution (asymptotic t-ratios in parentheses).

Model for low-pay probability conditional on	Low-pay		High-pay	
Assumption on initial conditions	Endogenous	Exogenous	Endogenous	Exogenous
experience/10	-0.0687	-0.2051	-0.0158	-0.0887
	-(0.6465)	-(2.9220)	-(0.3004)	-(1.6790)
education>=high school	-0.6398	-0.7816	-0.2286	-0.3117
	-(2.6898)	-(3.5180)	-(1.6147)	-(2.1380)
female	0.2423	0.4276	0.1978	0.3192
	(1.2268)	(2.6030)	(1.6943)	(2.7170)
non-manual	-0.1347	-0.2926	-0.4401	-0.5346
	-(0.5511)	-(1.2330)	-(2.9895)	-(3.5580)
firm size>=100	-0.4091	-0.6501	-0.3924	-0.5494
	-(1.4099)	-(2.5290)	-(2.7651)	-(3.8800)
public sector	0.0762	-0.3013	-0.5340	-0.6996
	(0.2285)	-(1.1520)	-(3.5323)	-(4.5840)
agriculture	0.0772	0.2359	0.1064	0.3986
	(0.2472)	(0.7560)	(0.3359)	(1.2080)
bank, insurance, transport& communication retail trade, personal &household serv	0.2572	0.3181	0.0006	0.0431
	(1.5111)	(1.8470)	(0.0043)	(0.2970)
living in the north	-0.3727	-0.4688	-0.1635	-0.1880
	-(2.2171)	-(2.9740)	-(1.6044)	-(1.7760)
constant	0.9524	0.7741	-0.9903	-0.5871
	(5.1484)	(4.2920)	-(5.8353)	-(3.7240)
rho	-0.4583		-0.6468	
	-(1.8010)		-(3.4950)	
n obs	2148	334	2148	1814
pseudor2	0.2725	0.1329	0.2725	0.1661
pvalue	0.0000	0.0000	0.0000	0.0000

Notes: reference categories are 20 (vs 30, see text) years of experience, education<high school, male, manual, firm size<100, manufacturing, residence in the centre, south or isles.

Table 4.6. Endogenous switching bivariate probit estimated marginal effects for the conditional low-pay probability: Hourly wages

Low-pay threshold	Bottom quintile		Third decile	
Conditioning starting state	low-pay	high-pay	low-pay	high-pay
experience/10	-0.0300	-0.0015	-0.0294	0.0011
	-0.0687	*-0.0158*	*-0.0768*	*0.0075*
	(-0.6465)	(-0.3004)	(-1.2172)	(0.1503)
education>=high school	-0.2676	-0.0210	-0.0953	-0.0392
	-0.6398	*-0.2286*	*-0.2351*	*-0.2491*
	(-2.6898)	(-1.6147)	(-1.3634)	(-1.9422)
female	0.1053	0.0186	0.1102	0.0492
	0.2423	*0.1978*	*0.2808*	*0.3041*
	(1.2268)	(1.6943)	(1.9274)	(2.7750)
non-manual	-0.0583	-0.0451	-0.1369	-0.0843
	-0.1347	*-0.4401*	*-0.3346*	*-0.4811*
	(-0.5511)	(-2.9895)	(-1.7419)	(-3.6099)
firm size>=100	-0.1706	-0.0288	-0.0404	-0.0455
	-0.4091	*-0.3924*	*-0.0997*	*-0.3495*
	(-1.4099)	(-2.7651)	(-0.5139)	(-2.4608)
public sector	0.0332	-0.0502	-0.0988	-0.0490
	0.0762	*-0.5340*	*-0.2403*	*-0.3150*
	(0.2285)	(-3.5323)	(-1.0579)	(-2.1475)
agriculture	0.0337	0.0103	-0.1531	0.0427
	0.0772	*0.1064*	*-0.3636*	*0.2395*
	(0.2472)	(0.3359)	(-1.3761)	(0.6517)
bank, insurance, transport& communication, retail trade personal &household serv	0.1120	0.0001	0.1792	0.0130
	0.2572	*0.0006*	*0.4781*	*0.0831*
	(1.5111)	(0.0043)	(3.0072)	(0.6029)
living in the north	-0.1610	-0.0143	-0.1269	-0.0131
	-0.3727	*-0.1635*	*-0.3201*	*-0.0879*
	(-2.2171)	(-1.6044)	(-2.5041)	(-0.9472)
constant	*0.9524*	*-0.9903*	*1.2583*	*-0.9037*
	(5.1484)	(-5.8353)	(7.8214)	(-4.6931)
rho	-0.4583	-0.6468	-0.4690	-0.5307
	(-1.8010)	(-3.4950)	(-2.8237)	(-3.2501)
n obs	2148		2148	
pseudor2	0.2725		0.258	
pmod	0.0000		0.0000	
phead	0.0665		0.2270	
psel	0.0001		0.0000	
Estimated state dependence	0.4938		0.6104	
True state dependence evaluated at the characteristics of the low-paid	0.3389		0.4161	
True state dependence evaluated at the characteristics of the high-paid	0.1939		0.3040	

Notes: estimated coefficients in italic, asymptotic t-ratios of ML coefficients in parentheses, phead is the p-value from a LR test for the exclusion of the instruments in the headline equation, psel is the p-value from a LR test for the inclusion of the instruments in the selection equation, pmod is the model's p-value, reference categories as in Tab. 4.5

Table 4.6 (continued). Endogenous switching bivariate probit estimated marginal effects for the conditional low-pay probability: Monthly wages

Low-pay threshold	Bottom quintile		Third decile	
Conditioning starting state	low-pay	high-pay	low-pay	high-pay
experience/10	-0.0081	-0.0092	-0.0001	0.0017
	-0.0185	*-0.0649*	*-0.0002*	*0.0120*
	(-0.1702)	(-1.3213)	(-0.0038)	(0.2519)
education>=high school	-0.0506	-0.0188	-0.1128	-0.0288
	-0.1153	*-0.1520*	*-0.2421*	*-0.1934*
	(-0.5029)	(-1.2026)	(-1.4731)	(-1.5130)
female	0.1512	0.0468	0.1901	0.0614
	0.3466	*0.3478*	*0.4148*	*0.3814*
	(1.6227)	(3.2913)	(2.8714)	(3.5271)
non-manual	-0.1960	-0.1010	-0.0294	-0.0765
	-0.4496	*-0.6855*	*-0.0633*	*-0.4595*
	(-1.7753)	(-5.1487)	(-0.3591)	(-3.4210)
firm size>=100	-0.1057	-0.0245	-0.0331	-0.0434
	-0.2406	*-0.2231*	*-0.0709*	*-0.3484*
	(-0.9243)	(-1.6194)	(-0.3790)	(-2.4851)
public sector	0.0746	0.0045	-0.0178	-0.0124
	0.1714	*0.0368*	*-0.0382*	*-0.0856*
	(0.6672)	(0.2743)	(-0.2196)	(-0.6314)
agriculture	-0.1585	0.1367	-0.0662	0.0976
	-0.3639	*0.6826*	*-0.1411*	*0.4790*
	(-1.0653)	(2.5969)	(-0.4896)	(1.6337)
bank, insurance, transport& communication, retail trade personal &household serv.	0.1254	0.0036	0.1015	0.0058
	0.2892	*0.0289*	*0.2223*	*0.0394*
	(1.6499)	(0.2124)	(1.5155)	(0.2818)
living in the north	-0.1997	-0.0107	-0.0635	0.0006
	-0.4597	*-0.0887*	*-0.1370*	*0.0043*
	(-2.7659)	(-0.9624)	(-1.1583)	(0.0468)
constant	*0.9767*	*-1.0374*	*0.8712*	*-1.2046*
	(5.4035)	(-6.6316)	(5.8356)	(-7.4571)
rho	-0.4718	-0.7233	-0.6095	-0.6747
	(-2.0983)	(-4.4092)	(-4.7347)	(-5.5251)
n.obs	2148		2148	
pseudor2	0.2334		0.2232	
pmod	0.0000		0.0000	
phead	0.3963		0.4415	
psel	0.0002		0.0000	
Estimated state dependence	0.5319		0.5567	
True state dependence evaluated at the characteristics of the low-paid	0.3771		0.3944	
True state dependence evaluated at the characteristics of the high-paid	0.2742		0.2860	

Notes: estimated coefficients in italic, asymptotic t-ratios of ML coefficients in parentheses, phead is the p-value from a LR test for the exclusion of the instruments in the headline equation, psel is the p-value from a LR test for the inclusion of the instruments in the selection equation, pmod is the model's p-value reference categories as in Tab. 4.5

Table 4.7. Comparison of marginal effects between binary and polychotomous specification of the transition equation in models for the probability of low-pay persistence; Hourly wages.

Low-pay threshold	Bottom quintile		Third decile	
Transition equation	Ordered	Binary	Ordered	Binary
experience/10	-0.0096	-0.0314	-0.0081	-0.0287
	-0.0209	*-0.0720*	*-0.0193*	*-0.0746*
	(-0.2592)	(-0.6802)	(-0.3460)	(-1.1814)
education>=high school	-0.1960	-0.2694	-0.0776	-0.0942
	-0.4389	*-0.6457*	*-0.1799*	*-0.2315*
	(-2.3251)	(-2.7225)	(-1.1768)	(-1.3428)
female	0.0723	0.1069	0.0717	0.1092
	0.1569	*0.2468*	*0.1698*	*0.2769*
	(0.9882)	(1.2525)	(1.3138)	(1.8992)
non-manual	-0.0401	-0.0598	-0.0977	-0.1364
	-0.0876	*-0.1386*	*-0.2253*	*-0.3320*
	(-0.4494)	(-0.5666)	(-1.3499)	(-1.7307)
firm size>=100	-0.0508	-0.1687	0.0135	-0.0365
	-0.1117	*-0.4051*	*0.0318*	*-0.0898*
	(-0.5258)	(-1.3826)	(0.1855)	(-0.4616)
public sector	0.0776	0.0296	-0.0324	-0.0958
	0.1670	*0.0681*	*-0.0753*	*-0.2321*
	(0.6506)	(0.2046)	(-0.3779)	(-1.0202)
agriculture	-0.1069	0.0356	-0.1814	-0.1551
	-0.2402	*0.0818*	*-0.4049*	*-0.3668*
	(-0.9154)	(0.2617)	(-1.7018)	(-1.3895)
bank, insurance, transport & communication, retail trade personal &household serv	0.1313	0.1121	0.2015	0.1813
	0.2842	*0.2582*	*0.5047*	*0.4814*
	(1.8887)	(1.5147)	(3.3637)	(3.0362)
living in the north	-0.1604	-0.1646	-0.1234	-0.1280
	-0.3511	*-0.3822*	*-0.2911*	*-0.3213*
	(-2.4687)	(-2.3002)	(-2.5152)	(-2.5201)
constant	*0.9310*	*0.9556*	*1.1624*	*1.2576*
	(5.8346)	(5.1153)	(7.8697)	(7.8328)
v0	*0.6011*		*0.2818*	
	(7.4652)		(6.5951)	
v1	*0.8414*		*0.7755*	
	(8.2928)		(9.7521)	
v2	*1.3376*		*1.0260*	
	(9.2229)		(10.4639)	
rho	-0.5824	-0.4509	-0.5658	-0.4766
	(-3.2373)	(-1.7691)	(-3.8483)	(-2.8700)
n obs	2148	2148	2148	2148
pseudor2	0.2667	0.3083	0.2612	0.2966
pmod	0.0000	0.0000	0.0000	0.0000
phead	0.7691	0.1579	0.3913	0.0966
psel	0.0001	0.0002	0.0000	0.0000

Notes: estimated coefficients in italic, asymptotic t-ratios of ML coefficients in parentheses, phead is the p-value from a LR test for the exclusion of the instruments in the headline equation, psel is the p-value from a LR test for the inclusion of the instruments in the selection equation, pmod is the model's p-value, reference categories as in Tab. 4.5

Table 4.7 (continued). Comparison of marginal effects between binary and polychotomous specification of the transition equation in models for the probability of low-pay persistence; Monthly wages.

Low-pay threshold	Bottom quintile		Third decile	
Transition equation	Ordered	Binary	Ordered	Binary
experience/10	0.0253	-0.0106	0.0134	0.0024
	0.0534	*-0.0243*	*0.0269*	*0.0051*
	(0.6471)	(-0.2213)	(0.4824)	(0.0797)
education>=high school	-0.0670	-0.0534	-0.1160	-0.1097
	-0.1418	*-0.1225*	*-0.2339*	*-0.2327*
	(-0.7299)	(-0.5318)	(-1.6148)	(-1.4205)
female	0.1163	0.1519	0.1414	0.1841
	0.2467	*0.3508*	*0.2879*	*0.3966*
	(1.4354)	(1.6217)	(2.2968)	(2.7268)
non-manual	-0.1192	-0.1966	0.0145	-0.0230
	-0.2529	*-0.4538*	*0.0294*	*-0.0489*
	(-1.2106)	(-1.7816)	(0.1919)	(-0.2777)
firm size>=100	-0.0332	-0.1062	0.0134	-0.0288
	-0.0702	*-0.2433*	*0.0272*	*-0.0610*
	(-0.3367)	(-0.9259)	(0.1694)	(-0.3279)
public sector	0.1121	0.0704	0.0153	-0.0130
	0.2387	*0.1628*	*0.0309*	*-0.0277*
	(1.1304)	(0.6295)	(0.2043)	(-0.1599)
agriculture	-0.2068	-0.1511	-0.1366	-0.0669
	-0.4541	*-0.3482*	*-0.2739*	*-0.1409*
	(-1.5217)	(-1.0169)	(-1.0503)	(-0.4913)
bank, insurance, transport & communication, retail trade personal &household serv	0.1435	0.1262	0.1352	0.1041
	0.3060	*0.2934*	*0.2794*	*0.2255*
	(1.9380)	(1.6704)	(2.0507)	(1.5451)
living in the north	-0.1473	-0.2004	-0.0585	-0.0644
	-0.3130	*-0.4645*	*-0.1183*	*-0.1371*
	(-2.2156)	(-2.7887)	(-1.1244)	(-1.1662)
constant	*0.8942*	*0.9757*	*0.8653*	*0.8734*
	(5.5909)	(5.3489)	(6.3075)	(5.8800)
v0	*0.3400*		*0.4527*	
	(6.0658)		(8.4924)	
v1	*0.7564*		*0.7373*	
	(8.1752)		(10.1046)	
v2	*0.9782*		*1.1638*	
	(8.8033)		(11.2813)	
rho	-0.5971	-0.4579	-0.6856	-0.6254
	(-3.8866)	(-2.0000)	(-6.3920)	(-4.8878)
n obs	2148	2148	2148	2148
pseudor2	0.2369	0.2691	0.2193	0.2513
pmod	0.0000	0.0000	0.0000	0.0000
phead	0.1708	0.4830	0.2058	0.3358
psel	0.0003	0.0006	0.0000	0.0000

Notes: estimated coefficients in italic, asymptotic t-ratios of ML coefficients in parentheses, phead is the p-value from a LR test for the exclusion of the instruments in the headline equation, psel is the p-value from a LR test for the inclusion of the instruments in the selection equation, pmod is the model's p-value, reference categories as in Tab. 4.5

Appendix A4: A comparison of the INPS and SHIW data

This Appendix analyses the properties of the wage variable across the two data sets utilised in this Thesis. The basic wage information is annual labour income in both cases. A major difference between the two sources which, due to the lack of information, cannot be controlled for in making this comparison is that while administrative INPS records are gross of income taxes, survey data from the SHIW are net of taxes. For this reason, statistics on raw wages will not be considered, and the attention will be focused on the effects of personal characteristics on wages. In both cases, data include overtime payments; in the SHIW case they also include the monetary value of fringe benefits. Another relevant difference is given by the available working time information; while INPS data report the number of weeks worked during the year, SHIW data report the number of months worked in the year and the number of hours worked on average per week. For this reason, the comparisons in this appendix are made in terms of annual figures; to control for heterogeneity in time supplied, monthly (SHIW) and weekly (INPS) wages are also considered. A further difference is given by the fact that SHIW data include the public and agriculture sectors. To increase the comparability of the two data sets, observations from both sectors have been discarded from the SHIW data. Comparisons are based on the 1995 cross-section in both cases; male and female data are analysed separately, with the age range restricted from 18 to 65.

Table A4.1 presents results from OLS regressions of logarithmic wages on a set of workers attributes. In each of the cases considered, wages present a concave profile with respect to age. For yearly labour incomes, the peak is located around the age of 45. After adjusting for time worked, the peak age slightly rises for SHIW male data (from 47 to 50.7), while for females it remains fixed at 43; differences are more evident in the INPS case (from 45 to 58 for males and from 46 to 57 for females). Estimated coefficients on occupational dummies are well determined and indicate differentials ranging from 21 to 45% for white collar workers and from 82 to 149% for managers, the reference category being blue collar workers. Wage differentials in favour of non-manual workers tend to be rather homogeneous across genders in the INPS data set; the differential for managers is instead higher for females in the SHIW case. Inter-industry wage differentials are next taken into account, the comparison category being given by the manufacturing sector. The direction of differentials for the construction sector is in accordance across the two data sets, being negative for males and non significative for females. The dummy for the transport and communication sector typically attracts a positive coefficient, with a negative coefficient (significant at the 15% level) characterising yearly male data in the INPS. The banking and insurance sector is characterised by a positive and significative coefficient in each case, with estimates from the INPS data which are nearly triple those from the SHIW. Patterns for the retail trade and hotel sector are less clear: in particular, while the differential estimated from the SHIW is negative for males and positive for females, estimates from the INPS are always negative. Wage differentials by firm size are well determined and indicate an increasing wage advantage as we move towards larger firms; the tendency for differentials estimated from the INPS to be larger is also present, but less evident than in previous cases, especially for females. Finally, coefficients on geographical dummies indicate higher wages in the north of the country, a smaller advantage accruing to the centre, both compared to the south and isles. It is worth stressing that while INPS data attribute the largest coefficient to the north-west, SHIW data suggest that wages are higher in the north-east; moreover, differences in the size of coefficients across the two data sets are not as evident as for the other effects considered. Looking at estimates from the SHIW it seems that regional wage differentials are more important for women.

A different perspective for the comparison of the two data sets is given in Table A4.2, where marginal effects from probit estimates of the probability of being low-paid in 1995 are reported. The low-pay threshold is the bottom quintile of the wage distribution computed from the samples used for the OLS regressions of Table A4.1.[69] The *ceteris paribus* low-pay probability presents an u-shape profile, in accordance with the OLS analysis; the minima of these profiles tend to occur at slightly lower ages than the maxima of the OLS regression, major differences characterising weekly INPS data. The patterns of occupational effects resemble OLS coefficients in that their absolute value is higher in INPS data; it can also be observed that these effects are stronger for females according to both data sets. Marginal effects for the construction sector tend to evolve in line with the OLS regression, being positive for males and non significative for females. The patterns of marginal effects for the other industries are less clear. The dummy for transports and communications attracts a negative effect in the SHIW male sample, while the opposite holds for INPS data; effects for females are typically non significative, the exception being a negative significative effect for female weekly wages in the INPS sample. Evidence for the banking and insurance dummy departs a bit from the OLS analysis, effects for male SHIW data being non significatively different from the manufacturing sector, while, in the rest of the cases, the effect is negative. For the retail trade and hotel sector, estimates are not well determined and tend to be negative, thus reverting, evidence from OLS regressions for the INPS data. Effects for the firm's size and the region are well determined and negative; in both cases they tend to be stronger for females in both data sets.

[69] Dummy variables happening to be "perfect classifiers" and observations scoring "1" on these dummies have been dropped from the analysis.

Table A4.1: OLS log-wage regressions (t-ratios)

SHIW	Males Yearly wages		Females Yearly wages		Males Monthly wages		Females Monthly wages	
age/10	0.697	(12.664)	0.766	(8.079)	0.477	(11.622)	0.513	(7.098)
age^2/100	-0.073	-(10.454)	-0.088	-(6.903)	-0.047	-(9.113)	-0.059	-(6.081)
white collar*	0.236	(11.841)	0.220	(7.315)	0.225	(15.125)	0.210	(9.162)
manager	0.853	(15.514)	1.243	(4.413)	0.827	(20.173)	1.197	(5.576)
construction	-0.104	-(3.985)	0.036	(0.344)	-0.037	-(1.876)	0.041	(0.512)
transport & communication	0.074	(2.230)	-0.017	-(0.194)	0.058	(2.325)	0.077	(1.149)
bank, insurance, financial and real estates services	0.079	(2.570)	0.105	(2.334)	0.103	(4.528)	0.063	(1.843)
retail trade & hotel	-0.023	-(0.906)	0.039	(1.149)	-0.019	-(1.017)	0.077	(2.998)
20<=firm size<100	0.102	(4.626)	0.098	(2.913)	0.082	(5.018)	0.129	(5.040)
100<=firm size<500	0.181	(7.006)	0.162	(3.984)	0.160	(8.291)	0.196	(6.333)
firm size >500	0.265	(11.325)	0.271	(6.293)	0.217	(12.478)	0.287	(8.746)
north-west	0.120	(5.479)	0.258	(6.076)	0.121	(7.373)	0.277	(8.550)
north east	0.156	(6.881)	0.266	(6.355)	0.140	(8.268)	0.285	(8.942)
centre	0.065	(2.724)	0.144	(3.303)	0.065	(3.686)	0.223	(6.740)
constant	8.207	(79.647)	7.836	(46.219)	6.245	(81.296)	5.888	(45.575)
adj-r2	0.428		0.304		0.506		0.371	
nobs	2277		929		2277		929	
INPS	**Males Yearly wages**		**Females Yearly wages**		**Males Weekly wages**		**Females Weekly wages**	
age/10	0.546	(18.797)	0.919	(14.137)	0.210	(13.078)	0.285	(9.463)
age^2/100	-0.060	-(18.328)	-0.098	-(12.716)	-0.018	-(10.107)	-0.025	-(7.023)
white collar*	0.399	(62.593)	0.456	(36.338)	0.359	(102.05)	0.340	(58.461)
manager	1.436	(87.863)	1.494	(17.331)	1.377	(152.56)	1.374	(34.445)
construction	-0.281	-(31.490)	-0.043	-(0.921)	-0.109	-(22.021)	-0.008	-(0.363)
transport & communication	-0.015	-(1.564)	0.064	(1.823)	0.014	(2.742)	0.077	(4.752)
bank, insurance, financial and real estates services	0.269	(26.827)	0.289	(15.067)	0.248	(44.759)	0.289	(32.505)
retail trade & hotel	-0.074	-(7.997)	-0.125	-(8.261)	-0.053	-(10.386)	-0.005	-(0.686)
20<=firm size<100	0.156	(20.497)	0.135	(9.178)	0.122	(29.076)	0.114	(16.768)
100<=firm size<500	0.246	(29.975)	0.200	(12.405)	0.193	(42.406)	0.194	(26.072)
firm size >500	0.357	(45.524)	0.298	(18.870)	0.274	(63.238)	0.259	(35.542)
north-west	0.229	(28.346)	0.248	(10.940)	0.107	(23.981)	0.057	(5.386)
north east	0.219	(24.546)	0.198	(8.359)	0.093	(18.793)	0.045	(4.077)
centre	0.193	(22.103)	0.183	(7.560)	0.077	(15.941)	0.025	(2.246)
constant	8.727	(138.89)	7.520	(56.523)	5.693	(163.97)	5.296	(86.007)
adj-r2	0.393		0.280		0.601		0.519	
nobs	42843		12727		42843		12727	

Table A4.2: Probit marginal effects for the probability of wages below or at the bottom quintile (t-ratios)

SHIW	Males Yearly wages		Females Yearly wages		Males Monthly wages		Females Monthly wages	
age/10	-0.365	-(9.410)	-0.751	-(6.360)	-0.315	-(8.600)	-0.707	-(5.840)
age^2/100	0.040	(7.920)	0.090	(5.680)	0.034	(7.060)	0.085	(5.230)
white collar*	-0.060	-(3.750)	-0.205	-(5.420)	-0.064	-(4.240)	-0.240	-(6.150)
manager	-0.074	-(1.610)						
construction	0.062	(3.200)	-0.043	-(0.310)	0.023	(1.380)	0.086	(0.600)
transport & communication	-0.059	-(2.220)	0.005	(0.040)	-0.062	-(2.570)	-0.065	-(0.550)
bank, insurance, financial and real estates services	0.027	(0.900)	-0.110	-(1.930)	-0.014	-(0.530)	-0.111	-(1.900)
retail trade & hotel	0.009	(0.530)	-0.044	-(1.080)	0.013	(0.810)	-0.117	-(2.830)
20<=firm size<100	-0.052	-(3.990)	-0.143	-(3.710)	-0.038	-(3.140)	-0.155	-(3.970)
100<=firm size<500	-0.058	-(3.630)	-0.170	-(3.710)	-0.067	-(4.670)	-0.246	-(5.440)
firm size >500	-0.098	-(6.150)	-0.248	-(4.900)	-0.100	-(6.520)	-0.300	-(5.980)
north-west	-0.101	-(7.610)	-0.328	-(7.200)	-0.084	-(6.770)	-0.342	-(7.190)
north east	-0.099	-(7.740)	-0.347	-(7.720)	-0.094	-(7.840)	-0.361	-(7.690)
centre	-0.064	-(4.760)	-0.243	-(5.220)	-0.053	-(4.180)	-0.229	-(4.600)
pseudor2	0.262		0.2159		0.2715		0.2426	
nobs	2277		927		2223		927	
INPS	**Males Yearly wages**		**Females Yearly wages**		**Males wages**		**Females Weekly wages**	
age/10	-0.283	-(18.230)	-0.724	-(13.800)	-0.168	-(13.650)	-0.526	-(9.550)
age^2/100	0.033	(18.430)	0.077	(12.330)	0.018	(12.740)	0.051	(7.830)
white collar*	-0.101	-(25.780)	-0.343	-(33.500)	-0.121	-(32.670)	-0.458	-(43.140)
manager	-0.111	-(11.880)						
construction	0.117	(21.450)	0.037	(0.940)	0.054	(13.310)	-0.021	-(0.510)
transport & communication	0.050	(8.540)	-0.036	-(1.110)	0.013	(2.740)	-0.078	-(2.090)
bank, insurance, financial and real estates services	-0.055	-(6.360)	-0.107	-(5.810)	-0.053	-(5.780)	-0.120	-(5.860)
retail trade & hotel	-0.003	-(0.570)	-0.036	-(2.950)	-0.013	-(3.270)	-0.106	-(8.680)
20<=firm size<100	-0.060	-(17.800)	-0.130	-(11.940)	-0.062	-(26.240)	-0.142	-(12.650)
100<=firm size<500	-0.088	-(24.330)	-0.202	-(18.090)	-0.085	-(34.360)	-0.256	-(23.530)
firm size >500	-0.148	-(37.240)	-0.261	-(22.580)	-0.155	-(46.380)	-0.344	-(29.720)
north-west	-0.114	-(27.650)	-0.199	-(10.940)	-0.087	-(26.510)	-0.123	-(6.350)
north east	-0.087	-(22.520)	-0.133	-(7.420)	-0.061	-(20.390)	-0.083	-(4.270)
centre	-0.071	-(17.980)	-0.083	-(4.420)	-0.039	-(12.120)	-0.024	-(1.150)
pseudor2	0.2045		0.2447		0.2745		0.3571	
nobs	42843		12679		41724		12679	

Chapter 5

Discontinuous wage profiles, endogenous selection and mobility: a simulated estimation approach

5.1 Introduction

In this Chapter the analysis of low-pay transitions will be extended to control for the potential bias induced by the discontinuity of wage profiles in the SHIW data. Estimation of the bivariate probit with endogenous switching in Chapter 4 is based on the sample for which a valid wage is observed at both ends of the transition investigated, while observations available only at the beginning or at the end of the transition are discarded from the analysis. Such a sample selection rule may lead to biased parameters' estimates if the propensity to have a valid wage observation in both of the SHIW waves considered is not randomly distributed across individuals, but is correlated with unobservables in the transition equation.

A possible reason for such discontinuity could be *panel attrition*, i.e. unavailability of sample respondents in some of the panel waves. Movements out from the wage distributions could also be determined by economic or demographic factors (e.g., for example, unemployment or retirement), which do not imply attrition from the sample of respondents. In what follows, the two causes of intermittency will be analysed altogether, and I will term the resulting sample selection process *attrition from the wage distribution*.[70]

The sign of the correlation between the propensity to have continuous wage observations and low-pay transition probabilities is not clear a priori. On the one hand, sample selection could be determined by demand side factors, with workers abandoning the sample of wage earners as a consequence of events such as layoffs. In such a case, individuals dropping out from the sample over time are likely to be characterised by a low degree of attachment to the labour market and their characteristics (both observed and unobserved by the researcher) would probably positively influence the propensity to persist in low-pay. The exclusion of these

[70] In fact, this seems to be the meaning with which the term *attrition* is used by Bingley et al. [1995].

observations from the analysis would then lead to underestimation of both aggregate low-pay persistence and the effect of observable characteristics on transition probabilities. At the other extreme, exits from the data set may be the result of supply side decisions which, had the sample unit been observed in subsequent time periods, would have generated mobility out of low-pay, which is instead not observed due to the inability to track the missing observation. In this occurrence attrition from the wage distribution would negatively covary with low-pay persistence, so that inferences based on the "balanced" sample would overestimate both aggregate persistence and the effect of observable characteristics on transition probabilities. The actual situation will probably result from the interplay of these two effects and in the analysis which follows attention will be focused on the net result.

The case of endogenous attrition is an example of what Verbeek and Nijman [1992] define as a *non-ignorable sample selection rule*: conducting inference on the selected sample is legitimate only if conditioning on the availability of observations doesn't alter the joint density of the variables under examination, and only in this case the selection rule may be deemed ignorable. As pointed out in this study, given a set of incomplete data, there are three strategies which could be pursued. Data may first of all be imputed, i.e. missing bits of information are replaced by their prediction based on the available sample. Alternatively, available observations could be weighted in some way, in order to reconstruct their relative importance to what it should have been in a random (i.e. non attrited) sample. Finally, a model based strategy can be pursued. In this case, the treatment of the missing data process is deferred to the analysis stage, where the probability of belonging to the available data set is modelled jointly with the economic relation of interest. In the case of non-ignorability of the sample selection rule, this last strategy is superior in that both imputation and weighting would need to be model based in order to be properly carried out.

Such a modelling approach to attrition characterises the few studies which address the problem in the context of panel data on earnings. Hausman and Wise [1979] were concerned with endogenous selection in wage equations estimated on a sample of participants in the Gary Income Maintenance Experiment, in particular with attrition of subsequent responses once the observational unit has already been part of the sample. They have two waves of data and their model consists of a "random effect" log-wage equation plus a probit for the probability of staying in sample in the second period (retention probability). Correlation in unobservables is allowed between retention and wages in the second period, which in turn implies correlation between retention and wages in the first year through the "random effect" in wages. With this setup they can derive Heckman's correction term for wages in both waves. They find that the extent of bias is limited in statistical significance, while its sign implies that high wage workers tend to drop out from the experiment: this is consistent with the fact that high wage individuals benefit less from the experiment and are thus more likely to abandon it. They also find that attrition is more severe in simple analyses of variance rather than in structural models and suggest that this occurrence arises from the fact that in the latter case the conditioning set already includes the factors determining attrition (this point is also noted by Verbeek and Nijiman [1992]). Keane et al. [1988] were interested in analysing self-selection over the business cycle and thus to investigate the issue of wage cyclicality when macroeconomic shocks don't hit workers at random. The framework is again that of Heckman's selectivity correction in "random effect" log-wage equations, extended to multiple time periods. They found that attrition significantly biased wages in a procyclical direction, suggesting that high wage workers exit the employed pool during downturns. The issue of attrition in the context of wage mobility modelling[71] was addressed in Bingley et al. [1995], who used a trivariate probit[72] model to tackle selectivity of both initial conditions and attrition. Their model consists of a probit for the probability of having a valid wage at both ends of the transition investigated, an ordered probit for the starting wage decile and an ordered probit for mobility (descending, absent or ascending). It has to be stressed that in this study the kind of attrition analysed includes both genuine attrition and exits from the wage distribution, as is the case in the present Chapter. Their results point towards a statistically significative impact of attrition from the wage distribution on mobility, with attrition probabilities positively correlated with upward mobility.

This chapter implements trivariate probits to augment the low-pay transition model of Chapter 4 with an attrition (from the wage distribution) equation; following Bingley et al., double selectivity is modelled as a sequential nesting process. The use of trivariate normal integrals, which are not commonly packaged in statistical software, poses a computational difficulty for the implementation of maximum likelihood estimation. Such a problem is tackled here by means of simulation techniques, in particular by implementing the GHK simulator within STATA's maximum likelihood routines: details on this qualifying aspect of the present Chapter are given in the Appendix.

[71] In the study of Stewart and Swaffield [1999], the impact of attrition on the bivariate probit estimates of low-pay persistence is investigated by amalgamating exits from the wage distribution together with persistence in low-pay, not moving up the wage distribution being the common factor between these two outcomes. They find that results are not dramatically different when compared to those obtained on the balanced sample.

[72] Computation of the trivariate integral is carried out using quadrature routines.

5.2 Discontinuous wage profiles in the SHIW data

Before moving on to the modelling stage, this section describes the features of attrition from the wage distribution in the Bank of Italy's data set. In this context, an important aspect of the sampling design is the distinction between panel and non-panel households, the first group corresponding to those households sampled in (at least) two consecutive waves. Assignment to this group is carried out in two steps. In the first step, which takes place at the date of the first wave's interview (1993 in our case), each household is asked whether or not it is willing to be re-interviewed in the subsequent wave. At the second step, which takes place previous to interviews for the subsequent wave, roughly 50% of those households available for re-interview are sampled to take part in the new wave. In 1993, 87% of interviewed households (7040 out of 8089 households) gave their availability for a new contact in 1995; of these, 47% were actually re-interviewed. A limited number of households (299) were also re-sampled among those answering NO or DON'T KNOW at the question on availability for future interviews. On the whole, of the 8089 households forming the 1993 wave of the SHIW, 3645 belong to the panel sub-group.

Moving from the household to the individual level and focusing on the group of full-time wage earners with valid wage observations aged between 18 and 65 in 1993, which is the sample implicitly deemed to be randomly selected in Chapter 4 when cross-sectional probit regressions for the probability of being low-paid were carried out, such a sampling design implies that of the 5708 valid observations (see Table 4.1), only 2734 belong to panel households, of which 2160 (see again Table 4.1) have a valid wage in both 1993 and 1995.

The sampling process just described suggests that some caution should be exerted when defining the control group for the sample selection analysis: a considerable number of cases exit the sample at random, i.e. from a decision of the survey builders, and not for economic or demographic reasons. It would clearly make no sense to include these observations in an analysis of the probability of staying in the sample.

Table 5.1: Transition probabilities into 1995 status for the sample of 1993 wage earners aged from 18 to 65 belonging to panel households (low pay defined as bottom quintile of hourly wage distribution)

1993 wage status	Low-pay	High-pay
1995 status		
low-pay	0.388	0.051
high-pay	0.304	0.761
missing wage; part-time	0.049	0.037
self employed	0.027	0.01
entrepreneur	0.012	0.01
unemployed	0.092	0.019
retired	0.022	0.068
other	0.004	0
housewife	0.027	0.002
not observed	0.076	0.043
Total obs	490	2244

The analysis of this chapter thus enlarges the estimation sample to include those observations which belong to a panel household in 1993 but don't have an observable wage in 1995: these are observations which could have potentially stayed in the sample of wage earners, but are not observed in the arrival wage distribution, either because they left the employed labour force or the household of origin. Potentially, also those belonging to households refusing to cooperate (and not actually re-sampled) could have been used to form the control group; however, the reasons behind the willingness to cooperate in the subsequent wave are not clear and it has been preferred not to include these cases (a total of 367 individuals) in the analysis.[73]

[73] Similarly, individuals belonging to panel households and with a valid wage only in 1995 (34 cases) are not included in the control group of the attrition equation.

In order to get an illustration of the kind of movements out from the wage distribution which determine the attrition process, Table 5.1 gives the destinations in 1995 for the sample of wage earners belonging to a panel household and aged between 18 and 65 in 1993, i.e. the estimation sample for the present Chapter. Focusing on the comparison between low- and high-paid in 1993, it can be seen how the low-paid are characterised by higher transition rates especially in the group of the unemployed, and, to a minor extent, in the housewives and the "not observed" (i.e. *genuine* attrition) classes. On the other hand, the high paid have higher transition rates into retirement.

An alternative illustration of the selection process is given in Table 5.2, which reports results from probit regressions for the probability of having a valid wage in both 1993 and 1995 on a set of personal characteristics: the event under investigation is persistence in the sample of valid wage earners, while explanatory variables are measured at the beginning of the transition. The estimation sample differs from the one considered in Table 5.1 due to the presence of missing values in some of the explanatory variables; the same remark applies to the sample used in estimating the model of the next section.

Column 1 considers the effect of the wage determinants used in the reduced form low-pay probits of Chapter 4 (i.e. the selection equations of the bivariate switching probit), but without controlling for parental background. We can observe that the probability of persisting in sample displays an inverted u-shaped profile in labour market experience, indicating a higher sample attachment towards the central part of the working career, with maximum probability approximately 18 years after the beginning of the first job. Education has a positive impact on such a probability, while being female reduces it by 4 percentage points. Among the other variables considered, while geographical location, occupation and sectoral affiliations (within the private sector) have no significant impact, affiliation to the public sector or employment in large firms positively influence retention probabilities on a scale between 5 and 10 percentage points.

Table 5.2: Probit estimates (marginal effects) for the probability of having a valid wage in 1993 and 1995 (asymptotic t-ratios in parentheses).

	1		2	
experience/10	0.219	(9.73)	0.146	(5.74)
(experience/10)^2	-0.059	(-11.30)	-0.044	(-7.77)
education>=high school	0.038	(1.73)	0.048	(2.20)
female	-0.043	(-2.46)	0.017	(0.64)
living in the north	-0.005	(-0.30)	0.008	(0.49)
non-manual	0.006	(0.29)	0.010	(0.46)
firm size>=100	0.049	(2.36)	0.047	(2.23)
public sector	0.104	(4.63)	0.097	(4.35)
agriculture	-0.038	(-0.74)	-0.046	(-0.89)
service sector	-0.010	(-0.46)	-0.009	(-0.41)
dependent children			0.076	(2.96)
dependent children*female			-0.096	(-2.16)
married			0.099	(3.22)
married*female			-0.048	(-1.22)
per capita equivalized household wealth (millions of lire)			-0.153	(-2.99)
n. obs	2716		2716	
pseudo r2	0.0792		0.0971	

Column 2 augments the same specification with some reservation wage indicators; these are dummies for the presence of dependent children (i.e. aged less than 14) in the household and for being married, interacted with the gender dummy, and the per capita equivalised household wealth.[74] The first two variables are assumed to potentially influence workers' motivation to participate in the labour market in different directions depending upon gender: while for males the presence of family responsibilities is supposed to raise incentives to participate, for females

[74] Compulsory education usually lasted until 13 in the years examined. The equivalising factor for the wealth indicator is the square root of the number of household members.

negative signs could arise from the social structuring of children and household care. On the other hand, wealth is assumed to raise the reservation wage irrespective of gender. We can first of all observe that, with the exception of the gender dummy, all the other effects are robust to the inclusion of reservation wage indicators. The reasons for the loss in size and significance of the gender dummy become clear as we move to consider the new variables included in the regression. In particular, considering the indicator for the presence of dependent children and its interaction with the gender dummy, we can see how these effects come with the expected sign: a male with a dependent children within the household has a higher probability (7.6%) of staying in sample than an otherwise identical worker, while for females this probability is 2% lower than for otherwise identical males without dependent children within the household.[75] Taking the effect of marriage into account, it is rather strong (10%) and positive for males, while for females it is less intense and statistically significant. Finally, the household wealth indicator displays the expected negative sign.[76]

The probit analysis above has shed some light on the factors influencing exits from the sample of wage earners in 1995; in the next paragraph I will build on this in trying to assess the potential bias induced by attrition from the wage distribution within the model of low-pay transitions.

---

[75] Recall that the sample is not selected with respect the position in the household, i.e. this evidence is based also on 357 sons and 226 daughters, plus 55 other relatives. Thus, the sample departs a bit from a stylised model of labour supply allocation between husband and wife. Here I'm implicitly assuming that the factors determining the allocation of family responsibilities in the case of husbands and wives also influence the decisions of working sons and daughters.

[76] In a separate non reported analysis, the effect of wealth have been differentiated by sex, finding that for females such effect is still negative, but greater that the male one (the p-value on the estimated coefficient is 0.17).

5.3 A sequentially nested trivariate probit model for the analysis of sample selection bias induced by discontinuous wage profiles

This paragraph describes the modelling approach adopted in order to assess to what extent results obtained in the previous chapter are plagued by the presence of sample selection bias. Is it correct to focus the analysis of wage transitions only on those individuals for whom a valid wage is available in each panel wave? Or, on the contrary, are these individuals systematically different from the ones dropping out from the survey, so that their selection for estimation is endogenous, thus biasing estimation results?

To provide an answer to these questions, the (structure of the) model in Chapter 4 has been extended to allow for a third dichotomic event which interacts with the ones previously considered (i.e. low-pay/high-pay at both ends of the transition investigated) in determining the likelihood of the data. The resulting set-up is a trivariate conditional probit model which allows for sequential nesting of the equation of interest.

The modelling of attrition is carried out expanding the model for low-pay persistence proposed by Stewart and Swaffield [1999], i.e., differently from the model of Chapter 4, the 1995 wage outcome is assumed to be observable only for the 1993 low-paid. The model is expanded by acknowledging that it can be actually estimated only using observations for which a wage is observable in 1993 and 1995. Let R_i be a dummy partitioning the sample of 1993 wage earners[77] depending upon their wage observability in 1995. Recalling the notation from Chapter 4, let us also define d_{it} as dummy variables for low-pay occurrence in year t.

[77] Recall from section 5.2. that these are wage earners belonging to panel households, this last state being assumed exogenous to low-pay transitions.

At the first nesting level, a probit is estimated for the probability of having a valid wage in both periods.[78] At the second nesting level, only those observations for which $R_i=1$ are utilised to estimate a probit for the probability of low-pay in 1993. Finally, the sample of the 1993 low-paid observed in both waves is used to estimate a probit equation for low-pay in 1995. Of course, the probabilities of the three events above are simultaneously estimated, i.e. for those observation for which the 1995 wage outcome is observed the estimated probability is $Prob(R_i=1, d_{i93}=1, d_{i95}=1)= Prob(d_{i95}=1|d_{i93}=1, R_i=1)Prob(d_{i93}=1|R_i=1)Prob(R_i=1)$. It is worth stressing that the multivariate normal density assumed allows for unrestricted correlation between the errors, thus allowing a proper assessment of potential endogeneity issues among the three events investigated.

More formally, let us assume that the propensity to stay in sample (retention propensity) is a latent variable R^*; when R^* overcomes an unobservable (possibly individual specific) threshold τ^*, observations remain in the sample of wage earners in both waves. R^* is assumed to be a function of observable characteristics, and we only observe a dummy indicator signalling whether or not $R^*>\tau^*$:

$$\begin{aligned} R_i^* &= x'_{Ri}\,\delta_R + v_i \\ R_i &= I(R_i^* > \tau_i^*) \\ v_i &\sim N(0,1) \end{aligned} \tag{5.1}$$

where x'_{Ri} contains the whole set of explanatory variables used in the model.

The second stage can be formalised according to the discussion in Chapter 4 and assuming partial observability of the 1993 low-pay outcome:

[78] Note that specification in terms of retention, rather than attrition, characterises the models both of Hausman and Wise [1979] and Bingley et al. [1995].

$$\begin{aligned}
&g(w_{i93}) = x'_i\,\delta + u_i \qquad if \qquad R_i = 1\\
&d_{i93} = I(g(w_{i93}) \le g(\lambda_{93}))\\
&u_i \sim N(0,1)
\end{aligned} \tag{5.2}$$

The headline equation of interest is a probit equation for the occurrence of low-pay in 1995 for which two sources of partial observability are assumed:

$$\begin{aligned}
&h_1(w_{i95}) = z_i'\eta_1 + \varepsilon_{1i} \qquad if \qquad R_i = 1 \qquad and \qquad d_{i93} = 1\\
&d_{i95} = I(h_1(w_{i95}) \le h_1(\lambda_{95}))\\
&\varepsilon_1 \sim N(0,1)
\end{aligned} \tag{5.3}$$

Assuming that error terms in the three equations are jointly distributed as a tri-variate normal

$$\begin{pmatrix} v_i \\ u_i \\ \varepsilon_{i1} \end{pmatrix} \sim N_3\left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & & \\ \rho_1 & 1 & \\ \rho_2 & \rho_3 & 1 \end{pmatrix}\right] \tag{5.4}$$

and that observations are iid, the log-likelihood function of the model may be written as[79]:

$$\begin{aligned}
\log L = \textstyle\sum_i \{ & R_i d_{i93} d_{i95} \log \Phi_3(x'_{Ri}\,\delta_R x'_i\,\beta, z'_i\,\gamma_1; \rho_1, \rho_2, \rho_3) + \\
& R_i d_{i93}(1-d_{i95}) \log \Phi_3(x'_{Ri}\,\delta_R, x'_i\,\beta, -z'_i\,\gamma_1; \rho_1, -\rho_2, -\rho_3) + \\
& R_i(1-d_{i93}) \log \Phi_2(x'_{Ri}\,\delta_R, -x'_i\,\beta; -\rho_1) + \\
& (1-R_i) \log \Phi(-x'_{Ri}\,\delta_R) \}.
\end{aligned} \tag{5.5}$$

[79] The way δ commutes into β and $\eta 1$ in $\gamma 1$ is explained is Chapter 4.

The structure of the model can be compared with those of previous studies. As far as attrition is concerned, this Chapter assumes that data availability at the start of the transition is exogenous, as in Hausman and Wise [1979] and differently from Bingley et al. [1995], who include also those workers with a valid wage only at the end of the transition in the control group of their retention probit. Moreover, as mentioned above, while Hausman and Wise analyse attrition from the sample of respondents, either Bingley et al. and this Chapter consider attrition from the sample of wage earners. Taking the modelling of wage dynamics into account, Hausman and Wise have no dynamics in their specification (and hence no initial conditions problem), the present Chapter conditions wage levels on their lags, while Bingley et al. condition wage changes on lagged wages. It has to be stressed that the Bingley et al.'s specification and the one in this Chapter are observationally equivalent when mobility from the tails of the wage distribution is considered.

As for the bivariate model, restrictions in the form of variables entering only x_R would be desirable. Here I assume that such variables are given by some of the reservation wage indicators included in Table 5.1, namely the dummy for the presence of dependent children in the household interacted with the gender dummy. Such variables have been chosen on the basis of a reduced form bivariate probit model in which R_i and d_{i93} have been conditioned on a general specification of x_{Ri}[80]; results from the reduced form show how these two variables do not enter the low-pay equation significantly[81]; it is then assumed that their effect on wages in both time periods only works through participation in 1995. The choice of such variables is also in line with previous studies of attrition bias in panel wage analysis, namely with Keane et al. [1988], who use the number of kids as an instrument in their

[80] The specification includes the whole set of variables used in column 2 of Table 5.1 plus the parental background indicators of Chapter 4.
[81] The p-value for these variables in the 1993 low-pay equation is .77 for the male dummy and .9 for the female one.

employment equation. As mentioned in the introduction to the Chapter, attrition may well result from demand side factors, and it could also be argued that such factors are more relevant at the lower end of the wage distribution, where monopsonistic behaviour is likely to characterise the labour market (see Green et al. [1996]). However, it is difficult to imagine demand side factors, among the available information, which do not enter the wage equation directly.

As mentioned in the introduction, we can see from (5.5) that the log-likelihood function involves the cumulative density function of the trivariate normal distribution, whose evaluation has been implemented via simulation estimation: the appendix to this Chapter illustrates the practical implementation of such an estimator.

5.4 Results

Results from the simulated maximum likelihood analysis are given in Table 5.3, which reports SML estimated coefficients and asymptotic t-ratios for the two nesting equations and the low-pay transition equation; the analysis is restricted to the low-pay threshold defined in terms of the bottom quintile of the hourly wage distribution, while, as a benchmark, the first column of the Table reports results obtained with a nested bivariate model which only controls for the endogeneity of initial conditions.[82] The simulated likelihood function is computed using 75 random draws from the truncated normal distributions of interest.[83]

Column 2 reports results from a general specification of the trivariate model. By comparing estimated coefficients with the ones in the first column, we can observe that differences are not remarkable, a fact which suggests that results from the

[82] Results in column 1 are taken from Table 7 in Chapter 4, where bivariate and ordered probit specifications of the low-pay persistence equation were compared.
[83] In the appendix the performance of the SML estimator at different choices of the number of draws is checked, showing how estimates are robust to such a choice.

previous Chapter are robust to the addition of the attrition equation. This is confirmed by taking the estimated error covariance matrix into account. We can observe a positive correlation between retention probability and low-pay in the starting year conditional on retention, a negative correlation between retention and low-pay persistence conditional on retention and a negative correlation between low-pay in the starting year and low-pay persistence, both conditional on retention. None of the correlation coefficients is significant at conventional levels and the more precisely estimated is the correlation between low-pay level and low-pay persistence, which also preserves the sign and size it had in the analysis of Chapter 4. The correlation between retention and low-pay in the starting year is positive, a result which also arises in Bingley et al. [1995]. The correlation between retention and low-pay persistence is instead negative, meaning that those staying in sample have a lower propensity to remain in low-pay. However, neither is significantly different from zero.

Recalling the discussion from the introduction to this Chapter, such a result seems to indicate that observations abandoning the sample correspond to "weaker" labour market participants, and thus that their exclusion from the analysis could lead us to overestimate the whole phenomenon of persistence. As stressed above, however, the extent of the bias is irrelevant from the viewpoint of statistical significance.

As a next step in the analysis, restrictions are tested on the general model: after all, the previous finding of statistical insignificance of attrition bias could arise from the fact the structure imposed on the data is too complex to be precisely estimated, so that before concluding in favour of the irrelevance of attrition it is worth checking whether or not the finding is also supported by restricted specifications of the model in column 2.

Table 5.3: Simulated maximum likelihood estimates (asymptotic t-ratios) of the sequentially nested trivariate probit. GHK simulator with 75 draws.

	(1):bivariate probit without attrition		(2): unrestricted trivariate probit		(3): $\rho(R_i,d_{i93})=\rho(R_i,d_{i95})^*\rho(d_{i93},d_{i95})$		(4): restricted conditioning set		(5): (3)&(4)	
Low-pay 1995										
experience/10	-0.072	(0.680)	-0.079	(0.531)	-0.057	(0.408)	0.120	(1.456)	0.122	(1.546)
edu>=high sch.	-0.646	(2.723)	-0.657	(2.464)	-0.627	(2.381)				
female	0.247	(1.253)	0.259	(1.055)	0.225	(0.955)	0.034	(0.262)	0.030	(0.234)
non-manual	-0.139	(0.567)	-0.133	(0.474)	-0.110	(0.395)				
size>=100	-0.405	(1.383)	-0.432	(1.318)	-0.399	(1.233)				
public sector	0.068	(0.205)	0.033	(0.088)	0.066	(0.183)				
agriculture	0.082	(0.262)	0.112	(0.350)	0.095	(0.303)				
service sector	0.258	(1.515)	0.264	(1.536)	0.260	(1.518)				
living in the north	-0.382	(2.300)	-0.392	(2.101)	-0.370	(1.980)				
constant	0.956	(5.115)	0.968	(2.488)	1.041	(3.319)	0.713	(2.269)	0.747	(2.764)
Low-pay 1993										
experience/10	-1.109	(7.988)	-1.057	(5.865)	-1.084	(6.071)	-0.963	(4.946)	-1.000	-(6.387)
edu>=high sch.	-0.357	(3.063)	-0.342	(2.855)	-0.351	(2.956)				
female	0.636	(6.914)	0.618	(6.258)	0.629	(6.543)	0.236	(3.042)	0.242	(3.192)
non-manual	-0.540	(4.494)	-0.533	(4.416)	-0.538	(4.472)				
size>=100	-0.668	(5.787)	-0.648	(5.307)	-0.658	(5.499)				
public sector	-1.010	(8.056)	-0.977	(6.908)	-0.995	(7.204)				
agriculture	0.546	(2.582)	0.530	(2.486)	0.539	(2.532)				
service sector	0.103	(0.974)	0.098	(0.927)	0.101	(0.956)				
living in the north	-0.159	(1.848)	-0.158	(1.835)	-0.157	(1.820)				
(exp./10)^2	0.202	(6.338)	0.186	(4.136)	0.194	(4.334)	0.175	(3.518)	0.185	(4.729)
father blue coll.	-0.089	(0.908)	-0.082	(0.827)	-0.086	(0.861)	0.160	(1.965)	0.160	(1.956)
father not empl.	0.264	(0.876)	0.277	(0.900)	0.276	(0.893)	0.717	(2.751)	0.716	(2.731)
father missing	0.277	(1.725)	0.263	(1.580)	0.266	(1.584)	0.506	(3.107)	0.520	(3.314)
mother blue coll.	0.152	(0.823)	0.154	(0.824)	0.148	(0.790)	0.187	(1.187)	0.187	(1.183)
mother not empl.	0.239	(1.587)	0.233	(1.542)	0.237	(1.567)	0.124	(0.981)	0.127	(1.001)
mother missing	0.521	(3.056)	0.472	(2.465)	0.503	(2.778)	0.565	(2.922)	0.599	(3.661)
father edu.>=hs	-0.386	(2.082)	-0.382	(2.058)	-0.388	(2.099)	-0.631	(3.801)	-0.643	-(3.975)
mother edu>=hs	-0.284	(1.191)	-0.281	(1.182)	-0.284	(1.194)	-0.383	(1.889)	-0.388	-(1.904)
constant	0.488	(2.325)	0.393	(1.428)	0.441	(1.625)	-0.543	(2.502)	-0.505	-(2.660)

Table 5.3: continued

Retention										
experience/10			0.461	(4.755)	0.464	(4.781)	0.509	(5.370)	0.511	(5.400)
edu>=high sch.			0.160	(1.944)	0.160	(1.936)				
female			-0.038	(0.510)	-0.037	(0.498)	0.028	(0.394)	0.028	(0.387)
non-manual			0.031	(0.366)	0.033	(0.385)				
size>=100			0.177	(2.152)	0.176	(2.148)				
public sector			0.365	(4.240)	0.364	(4.232)				
agriculture			-0.139	(0.772)	-0.144	(0.802)				
service sector			-0.030	(0.381)	-0.030	(0.377)				
living in the north			0.003	(0.044)	0.002	(0.040)				
(exp./10)^2			-0.151	(7.047)	-0.151	(7.092)	-0.161	(7.640)	-0.161	-(7.734)
dep. children			0.354	(3.884)	0.353	(3.863)	0.343	(3.834)	0.340	(3.819)
dep.chil.*female			-0.433	(3.240)	-0.438	(3.276)	-0.425	(3.285)	-0.428	-(3.308)
father blue coll.			0.114	(1.643)	0.113	(1.636)	0.036	(0.540)	0.035	(0.529)
father not empl.			0.306	(1.146)	0.307	(1.151)	0.242	(0.912)	0.244	(0.922)
father missing			-0.107	(0.852)	-0.116	(0.933)	-0.210	(1.677)	-0.218	-(1.778)
mother blue coll.			-0.057	(0.446)	-0.061	(0.480)	-0.081	(0.648)	-0.084	-(0.677)
mother not empl.			-0.108	(1.145)	-0.108	(1.153)	-0.089	(0.973)	-0.090	-(0.980)
mother missing			-0.513	(4.233)	-0.512	(4.221)	-0.561	(4.688)	-0.563	-(4.708)
father edu.>=hs			-0.002	(0.015)	-0.008	(0.070)	0.070	(0.630)	0.065	(0.594)
mother edu>=hs			-0.072	(0.540)	-0.071	(0.532)	-0.005	(0.040)	-0.005	-(0.037)
constant			0.531	(3.454)	0.533	(3.463)	0.764	(5.524)	0.768	(5.582)
r(Ri,di93)			0.193	(0.591)			0.325	(0.777)		
r(Ri,di95)			-0.062	(0.110)	-0.187	(0.377)	-0.357	(0.796)	-0.412	-(1.088)
r(di93,di95)	-0.451	(1.769)	-0.420	(1.008)	-0.485	(1.337)	-0.516	(2.271)	-0.523	-(2.449)
n.obs		2148		2716		2716		2716		2716
logLik		-800.65		-2053.78		-2053.87		-2263.69		-2263.7

A first restricted version of the model is proposed in column 3, where the null hypothesis that the correlation between retention and initial low-pay is the product of the other two correlation coefficients is tested: $\rho(R_i,d_{i93})=\rho(R_i,d_{i95})*\rho(d_{i93},d_{i95})$. The hypothesis means that the correlation between initial low-pay and retention only works through the combination of the correlation between retention and final low-pay ($\rho(R_i,d_{i95})$) and the correlation between initial and final low-pay, i.e. the individual effect in earnings ($\rho(d_{i93},d_{i95})$); apart from this combined effect, there's no direct correlation between R_i and d_{i93}. This hypothesis is adopted from the outset of the analysis by Hausman and Wise [1979]. As discussed in the previous Section, the sample selection process in their study is similar to the one of this Chapter, which makes the hypothesis worth testing. Also, the signs (but not the sizes) of the correlation coefficients in column 2 are in accordance with this hypothesis. Results are given in column 3. By first considering the maximised simulated likelihood and comparing it to the one of the unrestricted model in column 2 via a Likelihood Ratio test, we obtain a $\chi 2$ statistic of 0.18, which strongly supports the non-rejection of the null. The impact on the estimated parameters is negligible as far as the effect of explanatory variables is concerned. Taking the two remaining correlation coefficients into account, we can see that they both gain in size and precision, with the correlation between initial low-pay and low-pay persistence approaching conventional levels of statistical significance. This remark doesn't apply to the correlation coefficient between low-pay persistence and sample retention (the main object of the analysis in this Chapter), whose sign still indicates that retention is negatively correlated with low-pay persistence.

Results up to this point indicate that the extent of attrition bias is pretty weak, a finding which also arises in Hausman and Wise [1979]. As shown by those authors, this finding emerges in a "structural" model of earnings, i.e. where the conditioning set contains a set of explanatory variables deemed to "cause" earnings, and which are likely also to determine the attrition process. This is actually true in their study: in a simple variance decomposition analysis of earnings they obtain a significant effect of attrition on earnings. A natural question which arises is then whether or not the finding of non significant attrition bias in our model of low-pay persistence is also due to the features of the conditioning set. To provide an answer to such a question, column 4 of Table 5.3 further simplifies the model of low-pay persistence, excluding both demand and supply side factors appearing in the transition equation from the model; consistently, such variables are also excluded from the selection equations.[84] By comparing the maximised simulated likelihood function with the one from column 2 with a Likelihood Ratio test, we obtain a $\chi 2$ statistic of 419.82, which is well above the critical values of the $\chi 2$ distribution with 21 degrees of freedom (7 variables are excluded from each equation) at usual confidence levels, thus clearly rejecting the restriction imposed. We can notice how the exclusion of the set of explanatory variables brings labour market experience to the verge of statistical significance in the low-pay transition equation, although with a reverse sign which arises from the fact that we are not controlling for other factors, for example education. On the other hand, the coefficient on the gender dummy looses both size and significance. Focusing on the estimated error covariance matrix, we can observe a general rise in size and precision for each correlation coefficient, in particular for $\rho(d_{i93},d_{i95})$, which is now significantly different from zero at conventional levels. Gains in precision also characterise the estimate of $\rho(R_i,d_{i95})$, but not enough to conclude in favour of the relevance of attrition bias. Thus, some effect of attrition seems to be present in this less-structured specification, but both the fact that the restricted model is not

[84] The two explanatory variables left in the transition equation are the gender dummy and the linear term in labour market experience. The reason for leaving these variables in the model is to maintain a quadratic profile in experience for the low-pay selection equation, and a comparison category for the "instruments" of the retention selection equation.

supported by the data and the unsatisfactory precision characterising the attrition bias parameter even in this case clearly suggest that attrition can be deemed ignorable in this case.

A final test for the relevance of attrition bias is reported in column 5, which combines the restrictions of columns 3 and 4, i.e. $\rho(R_i,d_{i93})=\rho(R_i,d_{i95})^*\rho(d_{i93},d_{i95})$ with the exclusion of structural explanatory variables from the conditioning set. Again, these restrictions are clearly rejected at conventional levels (the unrestricted model is the one in column 2). Taking the attrition bias parameter into account, we can observe a further gain in size and precision, which is, however, not enough to conclude in favour of the relevance of attrition bias.[85] Thus even if the combination of a restricted covariance matrix and a restricted conditioning set was supported by the data, its effect on the attrition parameter would not lead us to reject the analysis of Chapter 4 for suffering from sample selection bias.

5.5 Summary and conclusions

This Chapter has investigated the extent to which results obtained in the analysis of low-pay transitions of Chapter 4 are biased by potentially endogenous exits from the wage distribution, i.e. by the possibly non-random propensity with which workers with a valid wage at the beginning of the transition observed leave the sample of wage earners during such a transition.

Focusing on this problem of sample selection involved some computational difficulties: controlling for attrition bias required expanding the bivariate probit framework of Chapter 4 to include a third limited dependent variable equation and hence the resulting likelihood function included trivariate normal integrals which are

[85] The p-value for this coefficient is now 0.28.

not packaged within statistical software. The problem has been tackled by implementing a simulated maximum likelihood estimator, in particular adopting the so-called GHK simulator, using STATA's maximum likelihood routine; details on the construction of the simulated estimator are given in the Appendix.

Results obtained point towards the ignorability of sample selection bias: various versions of a sequentially nested trivariate probit model in which the first stage controls for non-random attrition from the wage distribution have been estimated on the SHIW data and in no case does the parameter measuring the extent of attrition reach statistical significance at conventional levels. The maximum level of precision for this parameter has been reached by restricting both the error covariance matrix and the conditioning set, with this last restriction not supported by the data. The sign of the parameter indicates that persisting in sample and persisting in low-pay negatively covary; had this parameter been significantly different from zero, this would have meant that the exclusion of attrited observations lead us to underestimate the true extent of low-pay persistence.

The finding of irrelevant attrition mirrors previous results from Hausman and Wise [1979] in the context of structural models of earnings. Opposite conclusions have been obtained in models of wage mobility by Bingley et al. [1995].

A final word of caution has to be issued in order to correctly interpret the results of this Chapter. The sampling design of the SHIW panel is peculiar in that about half of workers observed in the 1993 wave leave the sample at random due to a decision of the survey builders, and such observations have not been used in the estimation of the trivariate probit described above. As a consequence, the control group, i.e. attrited observations, in the analysis is relatively small (the balanced sample is enlarged by 26%), so that it may be that it doesn't provide enough variability to capture the effect of attrition. Results presented have thus to be viewed as contingent on the peculiar sample structure, and confirmation of such findings should be pursued in the future on panel data sets with more conventional sample design.

Appendix A5: Simulated maximum likelihood estimation of the trivariate probit model

The sequentially nested trivariate probit model utilised in this Chapter requires evaluation of trivariate normal integrals, thus posing computational problems due to the fact that the function evaluating such integrals is not available among those usually packaged within commonly used econometric software. Moreover, multiple integrals are hardly tractable by usual linear numerical approximations such as those based on the Newton Raphson method, and produce unreliable results in terms of the goodness of approximation (Hajivassiliou and Ruud [1994]).

As an alternative to numerical approximations, simulation based inference has been developed in recent years (see Stern [1997] for a survey and Gourieroux and Monfort [1996] for an extensive presentation of simulation estimation techniques and its applications in various context; see also Börsch-Supan et al. [1992], Börsch-Supan and Hajivassiliou [1993] and Hajivassiliou and Ruud [1994] for applications of the GHK (Geweke-Hajivassiliou-Keane)-smooth recursive conditioning simulator in the context of ML estimation of limited dependent variable models). The basic idea of simulated maximum likelihood estimation (SML) is to replace the intractable bit of the likelihood function by its simulated counterpart. This appendix illustrates how this is practically done in the case of the trivariate probit model, and shows the implementation of the SML-GHK estimator using STATA's maximum likelihood routine. The illustration of the method is carried out in terms of a (complete) trivariate probit, i.e., when full observability of the three variables is assumed. The STATA codes written to implement the simulated estimator are shown for both the trivariate probit and the sequentially nested trivariate probit used for the analyses of this Chapter.

The model of interest is a (seemingly unrelated) trivariate probit:

$$
\begin{aligned}
&Y_{ij}^{*} = X_{ij}\beta_j + u_{ij}\\
&Y_{ij} = I(Y_{ij}^{*} > 0)\\
&j = 1,2,3\\
&i = 1,..,N\\
&\begin{pmatrix} u_{i1}\\ u_{i2}\\ u_{i3}\end{pmatrix} \sim \left[\begin{pmatrix}0\\0\\0\end{pmatrix},\begin{pmatrix}1 & & \\ \rho_{12} & 1 & \\ \rho_{13} & \rho_{23} & 1\end{pmatrix}\right]
\end{aligned}
\tag{A5.1}
$$

where I(A) is a dummy indicating whether or not A is true. Assuming observations are i.i.d., the log-likelihood function of the model is:

$$
\begin{aligned}
\log L(\beta_j,\rho_{jl};X_j,Y_j) = \textstyle\sum_i \log\{&\Phi_3[K_{i1}X_{i1}\beta_1,K_{i2}X_{i2}\beta_2,K_{i3}X_{i3}\beta_3;\\
&K_{i1}K_{i2}\rho_{12},K_{i1}K_{i3}\rho_{13},K_{i2}K_{i3}\rho_{23}]\}\\
j,l &= 1,2,3\\
K_{ij} &= 2Y_{ij} - 1
\end{aligned}
\tag{A5.2}
$$

which involves the trivariate standard normal cumulative density function Φ_3 and is hardly tractable with traditional numerical approximation.

The main intuition behind the GHK smooth recursive conditioning simulator is to exploit the definition of conditional distribution functions[86]:

$$
\begin{aligned}
&\Pr(u_1 \le X_1\beta_1, u_2 \le X_2\beta_2, u_3 \le X_3\beta_3) =\\
&\quad \Pr(u_3 \le X_3\beta_3 | u_2 \le X_2\beta_2, u_1 \le X_1\beta_1)\Pr(u_2 \le X_2\beta_2 | u_1 \le X_1\beta_1)\\
&\quad \Pr(u_1 \le X_1\beta_1)
\end{aligned}
\tag{A5.3}
$$

[86] I drop i indices for notational convenience.

and to replace the joint multivariate normal with the product of sequentially conditioned univariate normal distribution functions. The expression in (A5.3) involves conditioning upon unobservables: if some approximation for these conditional distributions can be found, then the likelihood function only requires evaluation of univariate integrals which is feasible within ordinary statistical packages.

Consider the Cholesky decomposition of the errors' covariance matrix:

$$E(uu')=\Sigma=Cee'C' \tag{A5.4}$$

where C is the lower triangular Cholesky factor of Σ and $e \sim N_3(0,I_3)$, from which it follows that:

$$\begin{aligned} u_1 &= c_{11}e_1 \\ u_2 &= c_{21}e_1 + c_{22}e_2 \\ u_3 &= c_{31}e_1 + c_{32}e_2 + c_{33}e_3 \end{aligned} \tag{A5.5}$$

where c_{ji} is the C element in position ji.

Thus, we can re-write (A5.3) as:

$$\begin{aligned} &\Pr(u_1 \le X_1\beta_1, u_2 \le X_2\beta_2, u_3 \le X_3\beta_3) = \\ &\quad \Pr(e_3 \le (X_3\beta_3 - c_{32}e_2^* - c_{31}e_1^*)/c_{33})\Pr(e_2 \le (X_2\beta_2 - c_{21}e_1^*)/c_{22}) \\ &\quad \Pr(e_1 \le X_1\beta_1/c_{11}) \end{aligned} \tag{A5.6}$$

where e^*_1 and e^*_2 come from standard normal distributions with upper truncation points at $X_1\beta_1/c_{11}$ and $(X_2\beta_2-c_{21}e^*_1)/c_{22}$ respectively, i.e. they satisfy the conditioning events in (A5.3). It is worth stressing that we are now working with uncorrelated errors (the vector e) and that correlation between the elements of the original vector of errors u has been transferred to the truncation points of the sequential conditioning via the Cholesky decomposition of Σ.

Evaluation of the probability in (A5.6) involves unobservable terms e^*_2 and e^*_1. Let us introduce R random draws of e^*_1 and e^*_2, i.e. random draws of e_1 and e_2 from upper truncated standard normals, with truncation points given above. The GHK simulator of (A5.6) is the arithmetic mean of the R probabilities we obtain for each of these draws:

$$\begin{aligned}\tilde{P}_{GHK} &= \frac{1}{R}\sum_{r=1}^{R} \{\Pr(e_3 \le (X_3\beta_3 - c_{32}\tilde{e}^r{}_2 - c_{31}\tilde{e}^r{}_1)/c_{33}) \\ &\Pr(e_2 \le (X_2\beta_2 - c_{21}\tilde{e}^r{}_1)/c_{22})\Pr(e_1 \le X_1\beta_1/c_{11})\} = \\ &\frac{1}{R}\sum_{r=1}^{R} \{\Phi(X_3\beta_3 - c_{32}\tilde{e}^r{}_2 - c_{31}\tilde{e}^r{}_1)/c_{33}) \\ &\Phi((X_2\beta_2 - c_{21}\tilde{e}^r{}_1)/c_{22})\Phi(X_1\beta_1/c_{11})\}\end{aligned} \qquad \text{(A5.7)}$$

where $\tilde{e}^q{}_j$ is the q-th draw for e^*_j. The SML estimator is then obtained by replacing the cumulative trivariate normal distributions in (A5.2) by their simulated counterparts from (A5.7). Note that the resulting maximand will be conditional on the set of draws: for computational stability it is then important that such draws do not change with the parameter values during optimization steps (Hajivassiliou [1997]).

The last thing which is to be explained is how to generate random variables from upper truncated normal distributions. Such variables can be obtained by exploiting random number generators on the unit interval available in statistical packages and the inversion formula given, among others, in Stern [1997]. First of all, let us consider the relationship between draws from the uniform distribution on the unit interval (say v) and the corresponding random draws (say z) from the standard normal distribution; such a relationship is given by:

$$z = \Phi^{-1}(v). \tag{A5.8}$$

Draws for (say) upper truncated standard normals can be similarly obtained by recalling that in this case $F(z)=\Phi(z)/\Phi(b)$ where F(.) indicates the cumulative density function of the truncated variable and b is the upper truncation point; replacing F(z) by the uniform on the unit interval and solving the expression for z we get:

$$z=\Phi^{-1}(v\Phi(b)). \tag{A5.9}$$

Börsch-Supan and Hajivassiliou [1993] highlight the key features of the GHK simulator in the context of multivariate normal LDV models:

- simulated probabilities are unbiased;
- such probabilities are bounded in the (0,1) interval;
- the simulator is a continuous and differentiable function of the model's parameters.

They also show that GHK is more efficient, in terms of variance of probabilities' estimates, than other simulators such as the acceptance-rejection or the Stern simulator. Note that unbiasedness of simulated probabilities doesn't translate into unbiasedness of the logs of such probabilities, which is what is needed to compute the log-likelihood function. However, such bias becomes negligible as the number of draws is raised with the sample size (Hajivassiliou [1997]).

The remaining part of this appendix reports the STATA codes written to implement the SML estimator with GHK simulator for the trivariate probit model, together with some robustness checks performed.

```
capture program drop myll /*TRIPROB.DO: SURE trivariate
probit*/
program define myll
      version 5.0
      local lnf "`1'"
      local I1 "`2'"          /*declares Xiβi */
      local I2 "`3'"
      local I3 "`4'"
      local r21 "`5'"         /*declares the rho's*/
      local r31 "`6'"
      local r32 "`7'"
      local y1: word 1 of $S_mldepn    /*declares I(yj>0)*/
      local y2: word 2 of $S_mldepn
      local y3: word 3 of $S_mldepn
      mat A=I(3)                    /*places rho's into a
                                    matrix*/
      local i=1
      while `i'<=3 {
            local j=`i'+1
            while `j'<=3 {
                  mat A[`j',`i']=`r`j'`i''
                  mat A[`i',`j']=`r`j'`i''
                  local j=`j'+1
            }
            local i=`i'+1
      }
      capture mat C=cholesky(A) /*takes Cholesky factor*/
      if _rc==506 { /*if A<0 (i.e. not p.d.) => no action
                              taken*/
            di "A<0"
      }
      local i=2  /*takes the elements of the Chol. factor*/
      while `i'<=3 {   /*to feed them through the logL*/
            local j=1
            while `j'<=`i' {
                  mat ccc=C[`i',`j']
                  local c`i'`j'=trace(ccc)
                  local j=`j'+1
            }
            local i=`i'+1
      }
      local d=1
      while `d'<=$dr{
/*$dr is the number of draws, z1 and z2 are the uniform random
draws already generated outside the program*/
            /*generates the truncated normals*/
            local d11`d' "invnorm(z1`d'*normprob(`I1'))"
            local d10`d' "invnorm(z1`d'*normprob(-`I1'))"
            local d211`d' "invnorm(z2`d'*normprob((`I2'-
                  `d11`d''*`c21')/`c22'))"
            local d200`d' "invnorm(z2`d'*normprob((-`I2'-
                  `d10`d''*`c21')/`c22'))"
```

```
                /*the event space is partitioned into 8 (=2^3)
                        components*/
                /*(1,1,1) */
                local sp1`d' /*
                */"normprob((`I3'-`c32'*`d211`d''-
                    `c31'*`d11`d'')/`c33')*normprob((`I2'-
                    `d11`d''*`c21')/`c22')*normprob(`I1')"
                /*(0,0,0)*/
                local sp2`d' /*
                */"normprob((-`I3'-`c32'*`d200`d''-
                    `c31'*`d10`d'')/`c33')*normprob((-`I2'-
                        `d10`d''*`c21')/`c22')*normprob(-`I1')"
                        /*(0,1,1)*/
                local sp3`d' "(binorm(`I2',`I3',`r32')-`sp1`d'')"
                /*(1,0,1)*/
                local sp4`d' "(binorm(`I1',`I3',`r31')-`sp1`d'')"
                /*(1,1,0)*/
                local sp5`d' "(binorm(`I1',`I2',`r21')-`sp1`d'')"
                /*(1,0,0)*/
                local sp6`d' "(binorm(-`I2',-`I3',`r32')-
                            `sp2`d'')"
                /*(0,1,0)*/
                local sp7`d' "(binorm(-`I1',-`I3',`r31')-
                            `sp2`d'')"
                /*(0,0,1)*/
                local sp8`d' "(binorm(-`I1',-`I2',`r21')-
                            `sp2`d'')"
                local d=`d'+1
        }
        quietly replace `lnf'=0
        local d=1
        while `d'<=$dr {
                quietly replace `lnf'=`lnf'+`sp1`d'' if
                    `y1'==1&`y2'==1&`y3'==1
                quietly replace `lnf'=`lnf'+`sp2`d'' if
                    `y1'==0&`y2'==0&`y3'==0
                quietly replace `lnf'=`lnf'+`sp3`d'' if
                    `y1'==0&`y2'==1&`y3'==1
                quietly replace `lnf'=`lnf'+`sp4`d'' if
                    `y1'==1&`y2'==0&`y3'==1
                quietly replace `lnf'=`lnf'+`sp5`d'' if
                    `y1'==1&`y2'==1&`y3'==0
                quietly replace `lnf'=`lnf'+`sp6`d'' if
                    `y1'==1&`y2'==0&`y3'==0
                quietly replace `lnf'=`lnf'+`sp7`d'' if
                    `y1'==0&`y2'==1&`y3'==0
                quietly replace `lnf'=`lnf'+`sp8`d'' if
                    `y1'==0&`y2'==0&`y3'==1
                local d=`d'+1
        }
        quietly replace `lnf' =ln((`lnf')/$dr)
end
```

```
/*Fragment by which the sequentially nested trivariate probit
differs from the SURE trivariate probit*/


   local d=1
   while `d'<=$dr{

         /*generates the truncated normals*/
         local d11`d' "invnorm(z1`d'*normprob(`I1'))"
         local d211`d' "invnorm(z2`d'*normprob((`I2'-
         `d11`d''*`c21')/`c22'))"

/*the trivariate normal is required only for obs with both y1=1
(not attrited) and y2=1 (low-pay in 1993*/

         /*(1,1,1)*/
         local sp1`d' /*
         */"normprob((`I3'-`c32'*`d211`d''-
         `c31'*`d11`d'')/`c33')*normprob((`I2'-
   `d11`d''*`c21')/`c22')*normprob(`I1')"

         /*(1,1,0)*/
         local sp2`d' "(binorm(`I1',`I2',`r21')-`sp1`d'')"

         quietly replace `lnf'=`lnf'+`sp1`d'' if
         `y1'==1&`y2'==1&`y3'==1
         quietly replace `lnf'=`lnf'+`sp2`d'' if
         `y1'==1&`y2'==1&`y3'==0

         if `d'==$dr{
               quietly    replace    `lnf'=ln((`lnf')/$dr)    if
`y1'==1&`y2'==1
         }

         local d=`d'+1
   }
/*bivariate normal for observations not attrited but not low-
paid in 1993*/
   quietly replace `lnf'=ln(binorm(`I1',-`I2',-`r21')) if
   `y1'==1&`y2'==0

/*univariate normal for attrited observations*/
   quietly replace `lnf'=ln(normprob(-`I1')) if `y1'==0


end
```

Table A5.1: Comparison of bivariate probit ML and SML (R=75) estimates (asymptotic standard errors)

	ML		SML	
Y1				
x11	-1.2057	(0.3394)	-1.1987	(0.3395)
x12	0.2212	(0.0900)	0.2200	(0.0901)
x13	0.1244	(0.2338)	0.1360	(0.2340)
x14	0.1025	(0.2869)	0.0898	(0.2866)
Y2				
x21	-0.3997	(0.3480)	-0.3945	(0.3479)
x22	0.0455	(0.0932)	0.0438	(0.0932)
x23	0.3155	(0.2191)	0.3231	(0.2192)
x24	-0.5707	(0.3047)	-0.5718	(0.3049)
rho	0.7450	(0.0901)	0.7530	(0.0870)
n.obs	200		200	
logLik	-146.97		-146.603	

Table A5.2: Comparison of SML (R=100) trivariate probit estimates (asymptotic standard errors) between LIMDEP and STATA

	LIMDEP		STATA	
Y1				
x11	-0.9230	(0.2954)	-0.9227	(0.2616)
x12	0.7476	(0.1245)	0.7474	(0.1224)
x13	-1.6734	(0.2555)	-1.6537	(0.2392)
Y2				
x21	0.1721	(0.2627)	0.1661	(0.2607)
x22	-1.2202	(0.3306)	-1.1926	(0.3172)
x23	0.2218	(0.0861)	0.2150	(0.0835)
Y3				
x31	0.7563	(0.2536)	0.7607	(0.2307)
x32	-0.3801	(0.1051)	-0.3834	(0.0966)
x33	0.5648	(0.2048)	0.5554	(0.1915)
rho12	-0.1127	(0.1810)	-0.0950	(0.1758)
rho13	0.0742	(0.1453)	0.0878	(0.1438)
rho23	-0.5409	(0.1081)	-0.5445	(0.1181)
n.obs	200		200	
logLik	-272.069		-272.145	

Table A5.3: Behaviour of the STATA's SML estimator by different choices of R (asymptotic standard errors)

	R=75		R=100		R=150	
Y1						
x11	-0.9207	(0.2610)	-0.9227	(0.2616)	-0.9241	(0.2620)
x12	0.7461	(0.1221)	0.7474	(0.1224)	0.7477	(0.1226)
x13	-1.6178	(0.2304)	-1.6537	(0.2392)	-1.6610	(0.2415)
Y2						
x21	0.1647	(0.2618)	0.1661	(0.2607)	0.1505	(0.2595)
x22	-1.1752	(0.3179)	-1.1926	(0.3172)	-1.1647	(0.3129)
x23	0.2095	(0.0837)	0.2150	(0.0835)	0.2088	(0.0826)
Y3						
x31	0.7642	(0.2311)	0.7607	(0.2307)	0.7571	(0.2305)
x32	-0.3864	(0.0966)	-0.3834	(0.0966)	-0.3825	(0.0965)
x33	0.5436	(0.1919)	0.5554	(0.1915)	0.5593	(0.1915)
rho12	-0.0920	(0.1710)	-0.0950	(0.1758)	-0.1243	(0.1717)
rho13	0.0765	(0.1306)	0.0878	(0.1438)	0.0676	(0.1435)
rho23	-0.5155	(0.1206)	-0.5445	(0.1181)	-0.5504	(0.1185)
n.obs	200		200		200	
logLik	-272.8304		-272.145		-272.2269	

The robustness checks of SML estimator built using STATA's maximum likelihood routines are quite interesting. Table A5.1 compares a simulated bivariate probit estimate against its exact (or, more correctly, numerically approximated) counterpart, which is packaged within STATA. The sample utilised consists of 200 observations randomly extracted from the SHIW data set. Either size and significance of estimated parameters are very close between the two estimators; this is also true for the maximum of the log-likelihood function.

Table A5.2 compares estimates between the SML trivariate probit built in STATA and the one available in LIMDEP 7.0, which also uses the GHK simulator. It is worth stressing that this packaged estimator wouldn't have been sufficient for the analyses of this chapter, given that the model utilised here allows for sequential nesting via partial observability of two of the variables in the model, while the LIMDEP estimator is a pure multivariate probit, i.e. with full observability of each variable. Again size of coefficients, their standard errors and the value of the maximised likelihood functions are very similar across models. Finally Table A5.3 checks the sensitivity of the SML estimator with respect to the number of random draws used to approximate the trivariate integral: as can be seen, there are only minor differences when moving from one choice to another, suggesting that R=75 is sufficient. On the whole, evidence from the three Tables is supportive of the estimator built for this chapter in STATA.

Concluding remarks

This Thesis has used Italian panel micro-data to investigate the intertemporal wage covariance structure and the extent of mobility at the bottom of the wage distribution. As stressed in Chapter 1, analyses of inequality based on cross-sectional data focus on wage differentials at one point in time, but are not informative about the degree of persistence of such differentials over individual careers or the extent of hierarchical mobility which reduces the impact of inequality over the life-cycle. Analyses of wage persistence and mobility are thus needed to gain a complete picture of the dynamics of the wage distribution, and this Thesis has pursued such a task following two distinct methodological approaches: the minimum distance estimation of variance components models of the wage covariance structure and the multivariate microeconometric analysis of low-wage transition probabilities.

The covariance structure analysis of wage dynamics has formed the object of Chapter 2, where variance components models are estimated by minimum distance on an unbalanced administrative panel of male wages covering 1974-88. A descriptive analysis of the raw covariance structure has shown how remarkable changes have characterised wage dispersion over this period, with a phase of compression of the distribution which stops in the early 1980s and is followed by a reopening of differentials, particularly marked in the second half of the 1980s. These dynamics have already been documented by the existing literature on the Italian wage distribution, where consensus has emerged in ascribing them, at least in part, to institutional developments of the wage setting framework, in particular to the evolution of the wage indexation system, which moved from the full egalitarianism of the late 1970s to the abolition of automatic wage compensations for inflation in the early 1990s. The econometric analysis has been centred around the random growth model of permanent wages, which allows assessment of the extent of convergence and mobility within the permanent wage distribution. Results indicate that the distribution of permanent wages converged over the period analysed, as would be predicted by human capital theories of wage dynamics. However, extension of the analysis within subsamples defined by workers' occupations has revealed how such a convergence is absent from white collar data, casting doubts on an interpretation of the overall convergence based on the human capital paradigm and suggesting that such an outcome could have been imparted by the egalitarian wage indexation system, which was mainly effective in compressing wage differentials between occupational groups. On the other hand the divergence of white collar permanent wages is in line with the use of individual wage premia which characterised wage policies for this group since the mid-1980s.

Further insights into the intertemporal male wage covariance structure are provided in Chapter 3, where a larger unbalanced panel referring to the 1979-1995 interval has been analysed. Estimates of the raw covariance structure have shown how wage inequality has been growing also over the end of the 1980s and the first half of the 1990s and how these trends have been paralleled by increases in wage autocorrelation and hierarchical immobility. Estimated variance components models are characterised by flexible time shifters on each wage component which allow us to assess the role played by permanent differentials and wage volatility in determining aggregate inequality dynamics without relying on specific functional form assumptions. It has been shown how both components contribute to overall dynamics, with a predominant impact of permanent differentials which account for roughly 80% of the overall growth in inequality, their increase being concentrated in the central part of the 1980s, when egalitarian wage policies started to be abolished; on the other hand, increases in wage instability are observed over the first half of the 1990s, a finding which is in line with the higher "flexibility" characterising the Italian labour market in recent years. Random growth estimates of the permanent wage indicate that individual profiles within the overall distribution diverge, thus supporting the institutional interpretation of convergence put forward in Chapter 2, the more recent data set being less influenced by the compressionary effects of the wage indexation system. Moreover the random growth specification has been compared with a random walk one and the data seem to favour the latter, pointing towards a context of high permanent wage persistence. The analysis has next developed by assessing the relationship between observable workers' attributes and the dynamics of covariance components, showing how wage differentials between occupations are the main force behind the growth of permanent inequality. A model which decomposes covariance structure parameters according to workers' occupations has then been proposed. Results show how permanent wages are characterised by random walk processes which are similar across occupations, while, on the other hand, transitory shocks are more concentrated and persistent for white collar workers. Estimates of time shifters for the permanent wage indicate the presence of positive differentials in favour of white collar workers since the second half of the 1980s, suggesting that the growth in overall permanent inequality originated within the wage distribution for this group, confirming the relevance of the use of individual wage premia mentioned above. However, the model indicates that these policies have inflated the white collar permanent wage distribution, generating permanent inequality both between occupations and within white collar workers, but without altering individual wage dynamics between the two groups or inducing differentials in permanent wage mobility.

The analysis of low-wage mobility has been assessed in Chapter 4 using survey panel data for the 1993-95 transition. At the aggregate level, the Chapter has shown how raw low-pay persistence involves a considerable state dependence, the probability of being low-paid in the arrival wage distribution being much higher for those who are already low-paid in the starting year. The econometric analysis has been based on a bivariate probit model with endogenous switching which allows us to control for the potential endogeneity of the initial conditions of the wage process by modelling workers' selection into starting wage classes. Parental background indicators have been used to identify the selection equation and their validity as instruments has been found to be supported by the data. Parameter estimates show that the hypothesis of initial conditions exogeneity can always be rejected at conventional significance levels. A comparison between estimates obtained under the two alternative hypotheses on the nature of initial conditions has shown that assuming exogeneity systematically leads to overstatement of both size and significance of the effects of workers' attributes on transition probabilities. Once allowance is made for endogeneity, no effect on transitions can be detected from labour market experience. On the other hand, while some influence on low-pay persistence arises from workers' gender, education, geographical location or affiliation to the service sector, non-manual occupations and jobs in large firms seem to help in avoiding drops into the low-pay area from the upper part of the distribution. Model estimates have also been utilised to assess the extent of true state dependence within aggregate persistence, showing that the past experience of low-pay *per se* has a considerable effect on the future occurrence of the phenomenon. Data limitations, and in particular the fact that it has been possible to analyse only one transition, suggest the need to extend this kind of analysis on other data, once, and if, they became available.

The robustness of the conclusions reached in Chapter 4 to the presence of endogenous attrition from the sample of wage earners has been analysed in Chapter 5 by augmenting the low-wage mobility model with a third equation which controls for the probability of belonging to the balanced sample, i.e. the one used in the estimation of the model in Chapter 4. The resulting set-up is thus a sequentially nested trivariate probit, whose estimation is complicated by the fact that evaluation of trivariate normal integrals is required, whose computation is not feasible with the linear approximation algorithms normally employed by statistical software packages. To overcome the problem, a simulated maximum likelihood estimator, with which the intractable bits of the likelihood function are replaced by their simulated counterpart, has been implemented. Results indicate that estimates in Chapter 4 are robust to this generalization of the model to allow for panel attrition and that the extent of the attrition bias is not statistically significant. Various restrictions have been tested on the parametric structure and the conditioning set of the model, and also these simplified versions indicated that the hypothesis of exogenous attrition cannot be rejected.

Results produced by this study reveal how widening wage differentials to a large extent permanently affected the evolution of individual wage careers, especially within the distribution of white collar workers, an outcome which accords with the use of wage policies aimed at re-establishing occupational differentials which developed after the era of wage egalitarianism. However, such policies don't seem to have affected the individual specific components of wage profiles, thus denoting a certain inability in remunerating the time-varying aspects of workers skills. Moreover, the increased wage instability detected in recent years implies that the impact of cross-sectional inequality over the life-cycle is reduced at the cost of an increase in the uncertainty of labour incomes, so that the net impact on workers' welfare could be negative. On the other hand, while observable workers' attributes have been found to have a limited impact on the probability of abandoning the bottom of the wage distribution, the occurrence of low-pay episodes seems to worsen, by itself, the future developments of wage careers, suggesting that it could translate into persistent poverty. At the same time, the considerable extent of pure state dependence within aggregate persistence suggests that policies aimed at coping with the low-pay issue should focus on forms of direct wage or income protection, rather than on programs aimed at modifying the attributes of the low-paid. This scenario thus indicates that the Italian growth of wage differentials should be carefully dealt with by policy makers and prompts future research on this subject.

Bibliographical references

Abowd , J.M. - Card, D. [1989]: "On the Covariance Structure of Earnings and Hours Changes", Econometrica, vol. 57, no. 2, pp. 411-445

Abraham, K.G. - Houseman, S.N. [1993]: "Earnings Inequality in Germany", NBER, working paper no. 4541.

Accornero, A. [1992]: "La Parabola del Sindacato", Il Mulino.

Altonji, J.G. - Segal, L.M. [1996]: "Small Sample Bias in GMM Estimation of Covariance Structures", Journal of Business & Economic Statistics, vol. 14, no. 3, pp. 353-367.

Arkes, J. [1998]: "Trends in Long-Run versus Cross-Section Earnings Inequality in the 1970s and 1980s", Review of Income and Wealth, series 44, no. 2, pp.199-213.

Asplund, R. - Bingley, P. - Westergård Nielsen, N. [1998]: "Wage Mobility of Low-Wage Earners in Denmark and Finland", in Low Pay and Earnings Mobility in Europe, Asplund et al. (eds.), Edward Elgar.

Atkinson, A.B. - Bourguignon, F. - and Morrisson, C. [1992]: "Empirical Studies of Earnings Mobility", Harwood Academic Press.

Atkinson, A.B. - Cowell, F. (eds) [1983]: "Panel Data on Incomes", ICERD occasional papers, 2.

Baker, M. [1997]: "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings", Journal of Labor Economics, vol. 15, no. 2, pp. 338-375.

Baker, M. - Solon, G. [1998]: "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records", University of Toronto, mimeo.

Bardasi, E. [1996]: "Differenziali Salariali Pubblico-Privato", Lavoro e Relazioni Industriali, n.3.

Bardone, L. - Gittleman, M. - Keese, M. [1998]: "Causes and Consequences of Earnings Inequality in OECD Countries", OECD, Paris, mimeo.

Bell, D.B. [1996]: "Skill-Biased Technical Change and wages: Evidence from a Longitudinal Data Set", The Labour Market Consequences of Technical and Structural Change, discussion paper series, no. 8.

Berman, E - Bound, J. - Machin, S. [1997]: "Implications of Skill-Biased Technological Change: International Evidence" CEP LSE, discussion paper no. 367.

Bertola, G. - Ichino, A. [1995]: "Crossing the River: a Comparative perspective on Italian Unemployment Dynamics", Economic Policy, vol. 21, pp. 359-420.

Bigard, A.- Guillotin, Y - Lucifora, C. [1998]: "An International Comparison of Earnings Mobility", Review of Income and Wealth, series 44, no.4, pp. 535-554.

Bingley, P. - Westergård Nielsen, N. [1995]: "The Feasibility on a Study on Earnings Mobility among OECD Countries", report for the OECD, August

Bingley, P. - Bjørn, J. - Westergård Nielsen, N. [1995]: "Determinants of Wage Mobility in Denmark 1980-1990", University of Aarhus, CLS working paper 10-95.

Blau, F.D. - Kahn, L.M. [1996]: "International Differences in Male Wage Inequality: Institutions versus Market Forces", Journal of Political Economy, 104 (4), pp. 791-836.

Björklund , A. [1993]: "A Comparison between Actual Distributions of Annual and Lifetime Income: Sweden, 1951-89", Review of Income and Wealth, series 39, no. 4. pp. 377-386.

Blundell, R. - Preston, I. [1998]: "Consumption Inequality and Income Uncertainty", Quarterly Journal of Economics, vol. 113, no. 2, pp. 603-640.

Borjas, G.J. - Ramey, V.A. [1994]: "Time-series Evidence on the Sources of Trends in Wage Inequality", American Economic Review, Papers and Proceedings, pp. 10-16.

Börsch Supan, A. - Hajivassiliou, V. - Kotlikoff, L. - Morris, J. [1992]: "Health, Children and Elderly Living Arrangements. A Multiperiod - Multinomial Probit Model with Unobserved Heterogeneity and Autocorrelated Errors", in Topics in the Economics of Ageing, D. Wise (ed.), The University of Chicago Press, pp. 79-104.

Börsch Supan, A. - Hajivassiliou, V. [1993]: "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models", Journal of Econometrics, 58 (3), pp.347-368.

Bound, J. - Johnson, G. [1992]: "Changes in the Structure of Wages in the 1980s: an Evaluation of Alternative Explanations", American Economic Review, 82, pp.371-392.

Bourguignon, F. - Morrisson, C. [1983]: "Earnings Mobility over the Life-Cycle: a 30 Years Panel of French Cadres", in Panel data on Incomes, Atkinson A.B. and Cowell F.A. (eds.), ICERD, occasional papers, no.2.

Brandolini, A. - Sestito, P. [1994]: "Cyclical and Trend Changes in Inequality in Italy, 1977-1991", Banca d'Italia, mimeo.

Cannari, L. - Gavosto, A. [1994]: "L'Indagine della Banca d'Italia sui Bilanci delle Famiglie: una Descrizione dei Dati sul Mercato del Lavoro", Economia e Lavoro, XXVIII, n.1, pp. 63-79.

Casavola, P. - Gavosto, A. - Sestito, P. [1995]: "Technical progress and wage dispersion in Italy: Evidence from Firms' Data", Annales d'Economie et Statistique, 41-42.

Chamberlain, G. [1984]: "Panel Data", in Handbook of Econometrics, vol.2, ch. 22, Griliches Z. and Intriligator M.D. (eds.), North-Holland.

Contini, B. - Filippi, M. - Villosio, C. [1998]: "Earnings Mobility in the Italian Economy", in Low Pay and Earnings Mobility in Europe, Asplund et al. (eds.), Edward Elgar.

Crépon, B. - Mairesse, J. [1997]: "The Chamberlain Approach", in The Econometrics of Panel Data, Mátyás L. and Sevestre P. (eds.), second revised edition, Kluwer Academic Publishers, pp. 323-391.

Dell'Aringa, C. - Lucifora, C. [1994]. "Wage Dispersion and Unionism: Do Unions Protect Low Pay?", International Journal of Manpower, vol.15, no.2-3, pp. 150-169.

Dickens, R. [1996]: "The Evolution of Individual Male Earnings in Great Britain: 1975-94", CEP LSE, discussion paper no. 306.

Dickens, R. [1997]: "Caught in a Trap? Wage Mobility in Great Britain: 1975-94", CEP LSE, discussion paper no. 365.

Erickson, C.L. - Ichino, A. [1995]: "Wage Differentials in Italy: Market forces, Institutions and Inflation", in Differences and Changes in the Wage Structure, Freeman R. and Katz L.F. (eds.), Chicago University Press.

Entorf, H. - Kramarz, F. [1997]: "Does Unmeasured Ability Explain the Higher Wages of New Technology Workers?", European Economic Review, vol. 41, no. 8, pp. 1489-1510.

Fortin, N. - Lemieux, T. [1997]: "Institutional Change and Rising Wage Inequality", Journal of Economics Perspectives, 11, pp.75-96.

Gosling, A. - Machin, S. [1995]: "Trade Unions and the Dispersion of Earnings in UK Establishments", Oxford Bulletin of Economics and Statistics, vol. 57, no. 2, pp. 167-184.

Gosling, A. - Machin, S. - Meghir, C. [1998]: "The Changing Distribution of Male Wages in the UK, 1966-1992", IFS, wp no.98/13.

Gottschalk, P. [1997]: "Inequality, Income Growth and Mobility", Journal of Economic Perspectives, 11, pp. 21-40.

Gottschalk, P. - Moffitt, R. [1994]: "The Growth of Earnings Instability in the U.S. Labour Market", Brookings Papers on Economic Activity, 2, pp.217-254.

Gottschalk, P. - Smeeding, T.M. [1997]: "Cross-National Comparisons of Earnings and Income Inequality", Journal of Economic Literature, XXXV, pp. 633-687.

Gourieroux, C. - Monfort, A [1996]: "Simulation-Based Econometric Methods", CORE Lectures, Oxford University Press.

Green, F. - Machin, S. - Manning, A. [1996]: "The Employer-Size Wage Effect: Can Dynamic Monopsony Provide an Explanation?", Oxford Economic Papers, vol. 48, no. 3, pp. 433-455.

Gregory, M - Elias, P. [1994]: "Earnings Transitions of the Low Paid in Britain, 1976-91: a Longitudinal Study", International Journal of Manpower, vol. 15, no. 2/3, pp. 170-188.

Gullotin, Y - Bigard, A. [1992]: La Mobilité Hiérarchique des Salaires en France de 1967 à 1982, Economie et Prévision, 102-103, pp.189-204.

Guillotin, Y-Hamouche, S. [1998]: "Capital Humain, Conjonctore Économique et Mobilité Salariale", Université du Maine, mimeo.

Hajivassiliou, V. - Ruud, P. [1994]: "Classical Estimation Methods for LDV Models Using Simulation", in Handbook of Econometrics, vol. IV, Engle and McFadden (eds.), pp. 2383-2441.

Hajivassiliou, V. [1997]: "Some Practical Issues in Maximum Simulated Likelihood", STICERD discussion paper no. EM/97/340.

Hart, P.E. [1983]: "The Size Mobility of Earnings", Oxford Bulletin of Economics and Statistics, 45, pp.181-194.

Hause, J.C. [1980]: "The Fine Structure of Earnings and the On-the-Job Training Hypothesis", Econometrica, 48, 4, pp.1013-1029.

Hausman, J. - Wise, D. [1979]: "Attrition Bias in Experimental Panel Data: The Gary Income Maintenance Experiment", Econometrica, 47, pp. 455-473.

Heckman, J.J. [1981a]: "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time - Discrete Data Stochastic Process", in Structural Analysis of Discrete Data with Economic Applications, Manski C.F. and McFadden D. (eds.), Mit Press.

Heckman, J.J. [1981b]: "Heterogeneity and State Dependence", in Studies in Labour Markets, Rosen S. (eds), Chicago University Press.

Jarvis, S. - Jenkins, S.P. [1997]: "Low Income Dynamics in 1990s Britain", Fiscal Studies, vol. 18, no. 2, pp. 123-142.

Jarvis, S. - Jenkins, S.P. [1998]: "How Much Income Mobility is there in Britain?", The Economic Journal, 108, pp. 428-443.

Johnson, G.E. [1997]: "Changes in Earnings Inequality: the role of Demand Shifts", Journal of Economic Perspectives, 11, pp. 41-54.

Juhn, C. - Murphy, K.M. - Pierce, B. [1993]: "Wage Inequality and the Rise in Return to Skill", Journal of Political Economy, vol. 101, no. 3, pp. 410-442.

Katz, L.F. - Murphy, K.M. [1992]: "Changes in Relative Wages, 1963-1987: Supply and Demand Factors", Quarterly Journal of Economics, February, pp. 35-78.

Keane, M - Moffit, R. - Runkle, D. [1988]: "Real Wages over the Business Cycle: Estimating the Impact of Heterogeneity with Micro-Data", Journal of political Economy, 96, pp.1232-1266.

Krueger, A.B. [1993]: "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989", Quarterly Journal of Economics, February, pp. 33-60.

Lee, L.F. [1978]: "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited Dependent Variables", International Economic Review, vol 19 no. 2, pp. 415-33.

Levy, F. - Murnane, R.J. [1992] "U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations", Journal of Economic Literature, vol. XXX, pp. 1333-1381.

Lillard, L.A. - Weiss, Y. [1979]: "Components of Variation in Panel Earnings Data: American Scientist 1960-70", Econometrica, Vol. 47, no. 2, pp. 437-454.

Lillard, L.A. - Willis, R.J. [1978]: "Dynamic Aspects of Earnings Mobility", Econometrica, vol. 46, no. 5, pp. 985-1012.

Lucifora, C. [1995]: "L'analisi del Mercato del Lavoro con micro-dati", Economia e Lavoro, vol. 29, no. 3, pp. 3-20.

Lucifora, C. [1998]: "Working Poors? An Analysis of Low Wage Employment in Italy", in Low Pay and Earnings Mobility in Europe, Asplund *et al.* (eds.), Edward Elgar

MaCurdy, T.M. [1982]: "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis", Journal of Econometrics, 18, pp.83-114.

Manacorda, M. [1997]: "Mind the Step. The Evolution of Male Wage Inequality in Italy and the Escalator Clause", CEP-LSE, discussion paper no. 862.

Meng, C. - Schmidt, P. [1985]: "On the Cost of Partial Observability in the Bivariate Probit Model", International Economic Review, vol. 26, no. 1, pp. 71-85

Moffitt, R. - Gottschalk, P. [1993]: "Trends in the Covariance Structure of Earnings in the US: 1969-1987", Brown University working paper 9/93

Murphy, K.M. - Welch, F. [1992]: "The Structure of Wages", Quarterly Journal of Economics, Ferbruary, pp. 284-326.

OECD [1993]: "Earnings Inequality: Changes in the 1980s" Employment Outlook, pp. 157-184.

OECD [1996]: "Earnings Inequality, Low-paid Employment and Earnings Mobility", Employment Outlook, pp. 59-108.

OECD [1997]: "Earnings Mobility: Taking a Longer Run View", Employment Outlook, pp.27 - 61.

O'Higgins, T. [1994]: "YTS, Employment and Sample Selection Bias", Oxford Economic Papers, 46. pp. 605-628.

Shorrocks, A.F. [1976]: "Income Mobility and the Markov Assumption" Economic Journal, 86, pp. 566-578.

Shorrocks, A.F. [1978a]: "Income Inequality and Income Mobility", Journal of Economic Theory, 19, pp. 376-393.

Shorrocks, A.F. [1978b]: "The Measurement of Mobility", Econometrica, vol. 46, no.5, pp. 1013-1024.

Sloane, P.J. - Theodossiou, I. [1996]: "Earnings Mobility, Family Income and Low-Pay", Economic Journal, 106, pp 657-666.

Smith, R.E. - Vavricheck, B. [1992]: "The Wage Mobility of Minimum Wage Workers", Industrial and Labour Relations Review, vol.46, no.1, p.82-88.

Stern, S. [1997]: "Simulation-Based Estimation", Journal of Economic Literature, vol. XXXV, pp. 2006-2039.

Stewart, M.B. [1999]: "Low Pay, No Pay Dynamics", in Persistent Poverty and Lifetime Inequality: The Evidence, Centre for Analysis of Social Exclusion, report no.5.

Stewart, M.B. - Swaffield, J.K. [1997]: "The Dynamics of Low Pay in Britain", in Job, Wages and Poverty: Patterns of Persistence and Mobility in the New Flexible Labour Market, P. Gregg (ed.).

Stewart, M.B. - Swaffield, J.K. [1999]: "Low Pay Dynamics and Transition Probabilities", Economica, 66, pp. 23-42.

Topel, R.H. [1994]: "Regional Labour Markets and the Determinants of Wage Inequality", American Economic Review, Papers and Proceedings, pp. 17-22.

Topel, R.H. [1997]: "Factor Proportions and Relative Wages: the Supply-Side Determinants of Wage Inequality", Journal of Economic Perspectives, 11, pp. 55-74.

Treu, T. (ed.) [1984]: "Gli Studi di Relazioni Industriali negli Anni Settanta", il Mulino.

Verbeek, M. - Nijman, T. [1992]: "Incomplete Panels and Selection Bias", in The Econometrics of Panel Data, Mátyás L. and Sevestre P. (eds.), Kluwer Academic Publishers, pp. 262-302.

Wood, A. [1994]: "North-South Trade, Employment and Inequality: Changing Fortunes in a Skill Driven World", Oxford, Clarendon Press.