UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Departamento de Fundamentos da Análise Económica

HETEROGENEITY AND DYNAMICS IN INDIVIDUAL WAGES AND LABOUR MARKET HISTORIES

by

Laura Hospido Quintana

Memoria presentada para optar al grado de
Doctor por la Universidad de Santiago de Compostela

Junio de 2007

Memoria presentada para optar ó grao de
Doutor pola Universidade de Santiago de Compostela

Xuño de 2007

HETEROGENEITY AND DYNAMICS IN INDIVIDUAL WAGES AND LABOUR MARKET HISTORIES

Esta memoria foi presentada o día 23 de Novembro de 2007 no Salón de Graos da Facultade de Económicas da Universidade de Santiago de Compostela, ante o Tribunal composto por:

Dr. Alberto Meixide Vecino (Presidente). Departamento de Fundamentos da Análise Económica. Facultade de CC. Económicas, Universidade de Santiago de Compostela.

Dr. Wenceslao González Manteiga (Secretario)**.** Departamento de Estatística e Investigación Operativa. Facultade de Matemáticas, Universidade de Santiago de Compostela.

Dr. Costas (Konstantinos) Meghir. Departamento de Economía, University College London.

Dr. Jean-Marc Robin. EUREQua - Equipe de Recherche en Economie Quantitative de Paris 1, Université de Paris 1-Panthéon Sorbonne.

Dr. Enrique Sentana Iváñez. Profesor de Econometría de Series Temporais no Centro de Estudios Monetarios y Financieros (CEMFI), Madrid.

Calificación: *Sobresaliente Cum Laude.*

To my family

Starvation is the characteristic of some people not *having* enough food to eat. It is not the characteristic of there *being* not enough food to eat.

AMARTYA SEN

La vejez empieza cuando se pierde la curiosidad.

JOSÉ SARAMAGO

A sociedade non pode en xustiza prohibir o exercicio honrado das súas facultades á metade do xénero humano.

CONCEPCIÓN ARENAL

Acknowledgements / Agradecimientos / Agradecementos

El proceso de elaboración de esta Tesis Doctoral ha contado con el apoyo de muchas personas. Estas líneas son para todas ellas.

En primer lugar quiero dar las gracias a Manuel Arellano por su excelente labor de supervisión, su paciencia y ayuda constantes. Me gustaría expresarle toda mi gratitud por la generosidad y entusiasmo con los que comparte su conocimiento de la economía en general y la econometría en particular. A él le debo mucho de lo aprendido en estos años de intenso trabajo y mi interés por esta rigurosa manera de medir los datos para tratar de entender el mundo en el que vivimos.

En deuxième lieu je tiens à remercier Stéphane Bonhomme pour ses commentaires et ses conseils précieux sur pratiquement tous les contenus de la thèse ainsi que pour son appui infatigable tout au long de ma recherche.

También quiero expresar mi agradecimiento por sus sugerencias a Samuel Bentolila, Jinyong Hahn, Pedro Mira y Enrique Sentana.

My stay at the Economics Department of University College London during the Fall of 2006 was a fantastic experience. I wish to thank Costas Meghir, Nicola Pavoni and Marcos Vera for comments and encouragement. Special thanks are due to Richard Blundell for his suggestions, help and time.

En el Centro de Estudios Monetarios y Financieros (CEMFI) he disfrutado de un entorno de investigación y humano muy estimulante. Debo agradecer al CEMFI como institución y a Rafael Repullo, su director, el haberme ofrecido estas excelentes condiciones de trabajo y haberme facilitado la asistencia a cursos y conferencias. Mi reconocimiento va también para todos los profesores y el personal del CEMFI.

Desexo agradecer ó Departamento de Fundamentos da Análise Económica da Universidade de Santiago de Compostela á miña admisión no seu programa de doutoramento. Agradezo a colaboración de Juan José Ares, Roberto Bande, Melchor Fernández, Alberto Meixide e, especialmente, do meu titor Carlos Ricoy pola confianza depositada en min. Gustaríame tamén agradecer a axuda de Salomé Rielo e Berta Lago nos aspectos burocráticos e administrativos do proceso.

No puedo olvidar mostrar mi agradecimiento por sus comentarios y apoyo a los doctorandos con los que he convivido estos años y, en especial: en Madrid, a Alicia Barroso, Cristian Bartolucci, Carlos González, Abel Elizalde, Juan Ramón García, Cristina López-Mayán, Alfredo Martín, David Martínez, Javier Mencía, Ildefonso Méndez, Ulrich Sauder y - sobre todo - Gema Zamarro, otro de los pilares de esta tesis; and in London, Tommaso Frattini, Aida Galiano, Josep Mestres and Carolina Ortega. I am also indebted to Mario Fiorini, Albrecht Glitz and Esteban Jaimovich for their help during the Job Market.

Durante estos últimos años he contado con la ayuda de muchos amigos. Desde aquí les envío mi agradecimiento a todos ellos en Madrid, Santiago, Sarria, Zaragoza, Coruña, Vigo, Barcelona y Londres.

Nunca me ha faltado el respaldo, ánimo y cariño de mi familia - mis padres, Almu y Ale - y de Iago, que siempre cuida de mí. Esta tesis les debe mucho y a ellos se la dedico.

Finalmente, deseo agradecer a la Fundación Caixa Galicia la financión de mis estudios de postgrado en el CEMFI y a la Fundación Ramón Areces la concesión de la Beca que ha hecho posible la realización de esta Tesis Doctoral.

Gracias. Merci. Thanks. Grazas.

Contents

List of Tables

List of Figures

Introduction

This doctoral thesis considers new models and estimation methods for the analysis of the wage distribution and the labour market histories, from a dynamic perspective. In this analysis I use panel data, that is, repeated observations over time for the same individuals.

It is a well-known fact that individual wages evolve over time. In the data, we observe different patterns due to the cyclical aggregate conditions of the economy. We also find heterogeneous wage profiles across groups of individuals according to different observed characteristics: gender, age, education, and many others. Lastly, even among quite homogeneous groups, there exists heterogeneity at the individual level (e.g. ability), unobserved to the econometrician, but which would have an impact over the evolution of earnings along the professional careers of workers.

Another well-established fact is that individuals move between different labour market states - they are alternately employed, unemployed or out of the labour force - and, conditioning on working, they transit between different jobs. The way how workers build their own work histories also differs across individuals and over time.

Therefore, the starting point of this thesis is the idea that differences in individual labour market histories may help to better understand differences on individual earnings dynamics[1]. For instance, in the case of gender differentials, we would expect that gender differences in work histories would help to explain a substantial part of the male-female wage gap. In fact, several arguments in the literature have connected job mobility with the

[1] Throughout, I use the terms earnings and wages indistinctly.

existence and persistence of the wage gap over time. It has been argued that if women job mobility is more restricted due to variables like husband's residence and children's care, then wage gains predicted by search and job-matching models (Burdett, 1978; Jovanovic, 1979) will be smaller (Keith and Williams, 1995). Similar arguments could be extended to heterogeneous individuals in other dimensions, either observed (for age, Topel and Ward (1992) documented a sizeable impact of mobility in earnings of young males) or unobserved, and even, they could be extended to heterogeneity at the individual-job specific level (Postel-Vinay and Robin (2002) stressed the relevance of match effects in a model with within and between jobs wage dynamics).

Specifically, this thesis deals with the consideration of different levels of heterogeneity, both observed (chapter 1) and unobserved (chapters 2 and 3), individual (chapter 2) and job-specific (chapter 3), in empirical models for the dynamics of the distribution of earnings and labour market trajectories of workers along their careers. Chapter 4 represents a technical contribution, useful in several economic applications.

The first chapter studies gender differences in the wage growth and job mobility of young workers using data from the Spanish section of the European Community Household Panel (1994-2001). First, I build an experience measure that - as opposed to the conventional potential experience variable - considers the existence of discontinuities in the professional career of workers and, second, I analyse job mobility patterns for males and females, separately. From the comparison between the proposed experience measure - accumulated experience - and the one used normally - potential experience - it turns out that wage returns to experience are higher with the more accurate measure and that difference is greater for women than men. This result suggests the existence of a gender wage penalty to interruptions. Regarding job changes, the findings indicate that turnover rates are similar for men and women among young workers. Differences come from the side of some characteristics that are relevant for one of the two groups and not for the other, specially in case of promotion or in transitions to non-employment. For men, holding a position with responsibility or having a family it turns out to be important when changing job. On the contrary, for women it is relevant the type of journey or the size of the firm. Finally, in addition to the gender penalty to interruptions, I also find that early-career wage growth is greater for men than for women, and this is specially true in years when job changes occur. Similar results have been documented for data from U.S. (Light and Ureta, 1992; Loprest, 1992), Italy (Del Bono and Vuri, 2006) and Finland (Napari, 2007).

The second chapter, main body of the thesis, contributes to the earnings dynamics literature modelling not only the unobserved individual heterogeneity and time series properties of the conditional mean of earnings given its past (as in Lillard and Willis, 1978; MaCurdy, 1982; Abowd and Card, 1989, among others), but also allowing for richer sources of heterogeneity and dynamics in the conditional variance (Meghir and Pistaferri, 2004). In particular, I propose a dynamic panel data model with individual effects both in the mean and in a conditional ARCH type variance function. The second contribution consists on shedding some light on how the volatilities of individual wages behave in a period of increasing aggregate inequality as it has happened in the last three decades in the U.S. (Juhn, Murphy, and Pierce, 1993).

From a methodological point of view, this chapter applies and extends new estimation methods based on corrected likelihood functions. The use of this newly developed bias-corrected likelihood approach makes it possible to reduce the estimation bias to a term of order $1/T^2$ in a fixed-T context. The small sample performance of bias corrected estimators is investigated in a Monte Carlo simulation study. The simulation results show that the bias of the maximum likelihood estimator is substantially corrected for designs that are broadly calibrated to the Panel Study of Income Dynamics.

The empirical analysis is conducted on data drawn from the 1968-1993 PSID. I find that it is important to account for individual unobserved heterogeneity and dynamics in the variance, and that the latter is driven by job mobility. I also find that the model explains the non-normality observed in logwage data. In the last part of the empirical analysis, I look at the model's implications for consumption growth, in a simple precautionary savings framework (Browning and Lusardi, 1996). The main result is that an increase in individual risk induces a reduction in current consumption, and this effect is more important for the less educated people, slightly significant for the graduate and insignificant for the college educated. This result goes in line with the idea that there are more insurance possibilities for these latter (Blundell, Pistaferri y Preston, 2005).

Directly connected with chapter 2, the third chapter develops a model that explicitly considers job changes in the dynamics of wages and in the heterogeneity pattern. I propose an error components model designed to more thoroughly describe the impact of job mobility on the dynamics and heterogeneity of individual wages than previous references. In particular, the specification proposed has two different parameters to capture dynamics within jobs and across jobs, and the unobserved heterogeneity shows a richer pattern, as well, composed of both individual and job-specific effects. The potential endogeneity of job mobility in relation to earnings is circumvented using an instrument variable estimation method that controls for those unobserved heterogeneity components.

In the empirical application, I use data on work histories drawn from the PSID, which allows the distinction between voluntary and involuntary job-to-job changes. With respect to the main results, I find that - once we control for individual and job-specific effects - the dynamics within jobs is almost zero, whereas across jobs is significant but small. For the dynamics, the distinction between voluntary and involuntary transitions turns out to be irrelevant. However, that distinction matters in the case of the components of the cross-sectional variance. The estimated variance of the job-specific effects represents around one third of the variance for the individual fixed effects. If I consider a subsample that only includes involuntary job changes, the estimated variance of the heterogeneity across jobs increases up to one half.

Finally, the fourth chapter represents a technical contribution related to the computational calculation in practice of bias corrections of the type presented in the second chapter. Chapter 4 considers estimation of non-linear panel data models that include multiple individual fixed effects. Estimation of these models is complicated both by the difficulty of estimating models with possibly thousands of coefficients and also by the Incidental Parameters problem (Neyman and Scott, 1948), that is, noisy estimates of the fixed effects when the time dimension is short contaminate the estimates of the common parameters due to the nonlinearity of the problem. The chapter shows how to use an iterated algorithm which simplifies estimation in a nonlinear model with multiple fixed effects and also discusses the application of this computational simplification to bias corrected concentrated likelihoods. Some Monte Carlo experiments illustrate the results.

Chapter 1

Gender differences in Wage Growth and Job Mobility of Young Workers in Spain[1]

1.1 Introduction

Gender labour market differentials have always existed. Wage differences are the most noticeable and therefore the most studied, but there are also gender differentials in participation rates, unemployment rates, job mobility, ...

It is well-establish that the gender wage gap grows with the age of the individuals, but it also grows among the young. Evidence for the US shows that, even in the first years of professional career, wage growth is smaller for women than for men. In particular, for the first four years of the professional career, men accumulate a wage growth of 36 per cent *versus* a 29 per cent in the case of women (Loprest, 1992). In the case of Italy, wages increases by 21 percent for men and 20.4 per cent for women three years after labour market entry, but the gap widens rapidly over time (Del Bono and Vuri, 2006). The relevance of this difference is enlarged by the fact that this is indeed the life period in

[1]This chapter is partly based on my Master's Thesis at the Centro de Estudios Monetarios y Financieros (CEMFI).

which individuals achieve larger wage increases[2].

Traditionally, male-female wage differentials are decomposed following the method proposed by Oaxaca (1973)[3] into one component due to differences in socio-economic characteristics and another component, that remains unexplained, so-called discrimination. This decomposition confronts comparable men and women in a given point in time, but ignores dynamic aspects of the careers. Therefore, the starting point of this work is the idea that gender differences in labour market histories may help to better understand a substantial part of the male-female wage gap.

This chapter studies gender differences in two main features of young workers' labour market histories: interruptions and job changes. The analysis is focused on the case of young workers, for whom it turns out feasible to construct complete labour market histories with the information available in the European Community Household Panel (ECHP, in ahead).

With regard to interruptions, I propose the construction of an experience measure that takes into account discontinuities in the labour market profiles. To my knowledge, this type of measure has not been used in empirical works for Spanish data (for US see Corcoran and Duncan, 1979; Sandell and Shapiro, 1980; Mincer and Ofek, 1982; Light and Ureta, 1995[4]), so this represents the first contribution of this study.

Secondly, I investigate job mobility patterns in the first years of professional career: which types of job changes are more likely for men and women, which factors have a significant influence on those transitions and how much do their wages vary when changing job. This analysis is similar to Booth and Francesconi (2000) but, contrary to these authors

[2]Murphy and Welch (1990) indicate that two thirds of the wage growth that a worker accumulates throughout his working life concentrate in the first ten years.

[3]Or generalizations of this method like Brown *et al.* (1980) and Neumark (1988).

[4]For U.S. data, some authors have implemented even more ambitious specifications that consider not only interruptions but also the moment when they take place. In my case, the practical implementation of this idea was not possible due to limitations in the data.

who only consider job to job changes, the second contribution of my analysis consists on considering a model of transitions from job to job and also from job to nonemployment[5]. In fact, this distinction turns out to be important because it is just in those transitions to nonemployment where most of the gender differences arise.

It is well-known that female workers have a lower attachment to the labour force than men, with potentially important consequences for human capital accumulation, experience accumulation and job mobility. Several arguments in the literature connect those aspects with the existence and persistence of the wage gap over time:

- It has been argued that the discontinuity in women's labor market attachment may reduced their investment in human capital and thus their wages (Mincer and Polachek, 1974; Corcoran and Duncan, 1979; Gronau, 1988).
- If employers expect women to stay less in their jobs than men, firms will be less willing to invest in their training resulting in lower rates of pay (Donohue, 1988; Sicherman, 1996).
- If women job mobility is more restricted due to variables like husband's residence, and children's care, wage gains predicted by job-matching and search models [6] (Johnson, 1978; Burdett, 1978; Jovanovic, 1979) will be smaller (Keith and Williams, 1995).
- Women who face major family responsibilities are more likely to make a major adjustment in their labour market hours also when they change jobs. To the extent

[5]Apart from using UK data, another difference is that their study is not focused on the youth.

[6]In a matching model, job mobility is the consequence of a voluntary change to a better position where the worker is more productive and receives a higher pay. Search models are based in the existence of imperfect information. In these models, jobs are experience goods. As time goes by, the firm acquires more information and it can adjust the salary better. Under this approach, job mobility is the result of a 'poor' matching looking for a better chance.

that wages are less important in the decision to change for women than men, this may lead to lower wages over their career (Altonji and Paxson, 1992).

- Firms may assign workers with higher turnover rates into occupations that require less skills and with lower capital intensity (Barron *et al.*, 1993). These factors, together with the lower market value of previous experience for women, may contribute to the existence of a gender wage gap.

In addition, Topel and Ward (1992) claim that job mobility is a key factor on the wage growth of young workers. Consequently, differences in the labor mobility of young men and women may contribute to explain differences in their wage growth.

In Spain, the first empirical studies that consider gender wage differentials are Moltó (1984) and Peinado (1990), but for very small samples. Later, the availability of micro data and the development of the methodology associated to wage decompositions, pushed the diffusion of studies like De la Rica and Ugidos (1995) or Hernández (1995)[7]. More recent works propose the analysis of these differences throughout all the distribution of wages (García *et al.*, 2001; Dolado and Llorens, 2004; Gardeazabal and Ugidos, 2005).

On the contrary, references on gender differences in job mobility are scarce. Again in Spain, we only found the work of García-Crespo (2001) on gender differences in the promotions, and the one of Caparrós *et al.* (2004) on mobility and wage discrimination[8].

The rest of the chapter is developed as follows. Section 1.2 describes the data used. Section 1.3 explains the experience measure proposed in this study as well as the results from the wage regressions. Section 1.4 analyses gender differences in job transitions, while the final section concludes.

---

[7]The two works estimate more general wage decompositions to consider the problem of self-selection of the sample. They use data from the Encuesta de Estructura de Conciencia y Biografía de Clase (1985) and the Encuesta sobre Discriminación Salarial de la Mujer (1987), respectively.

[8]García-Crespo uses data of the Encuesta de Estructura, Conciencia y Biografía de Clase (1991) and Caparrós et al. the Spanish section of the European Community Household Panel (1994-1997).

1.2 The Data

The data base used in this work is the ECHP, a longitudinal annual survey, designed and coordinated by Eurostat, that includes homogenous information among countries on income, employment status, types of job changes, calendar of activities, education, health, and other demographic characteristics.

I use data from the Spanish section throughout eight annual waves (1994-2001). Since the objective is to construct complete labour market histories, the study is restricted to individuals from 16 to 39 years that, after finalizing their studies, are observable in their first job. In addition, I require that if they work they would do it for more than 15 hours per week. I exclude individuals with spells of self-employment. Finally, I have a sample with 3263 observations, corresponding to 543 men and 577 women. Sample selection filters are described in Appendix 1.A.

Table A.1 shows the main descriptive statistics[9]. In the sample, few individuals are married (6 per cent of the men and 21 per cent of the women) and even fewer have children (2 per cent and 9 per cent, respectively) mainly due to the fact that average age in the sample is around 25 years. Males accumulate more experience and tenure (in the case of potential experience the discrepancy is caused by the fact that, in average, women in the sample are older), work more hours per week (40 *versus* 37 hours), receive a 5 per cent higher wages than women and they work in occupations related to college degrees or agriculture, fishing and manufacture. Females, on the other hand, are more educated (in the sample, half of the women have a college degree), they also work in occupations related to college degrees or administration and work more often part-time (25 per cent of the women *versus* 12 per cent of the men). In the sample, the proportion of fixed-term contracts is around 40 per cent for both groups.

---

[9]Tables and figures are enclosed at the end of the document.

1.3 Accumulated Experience

1.3.1 Building the Experience Variable

The experience measure called **potential experience,** $POTEX_{it}$, defined for a given individual i and time t as age$_{it}$ - years of schooling$_i$ - 6, implies an unlikely assumption. The assumption implicit under this measure is that individuals work continuously since they finish their studies. This implies, for instance, that two individuals with the same years of education begin to work at the same time and do not suffer any interruption from that moment. In practice, potential experience may approximate quite wrong the capacities acquired by different individuals throughout their professional career.

This problem is specially worrisome in the case of female earnings, since women seem more willing to interrupt their careers due to family matters, like care giving activities, both to children and elderly parents. But discontinuities can be also common among young male workers, due to periods of job-shopping or fixed-term contract endings (the second motive is specially relevant among young workers in Spain nowadays[10]).

As an illustration, I calculate the fraction of time that individuals of a subsample of the ECHP spend working in five years[11]. Table A.2 shows the proportion of males and females that work at least a given number of months over that period.

We can see that being continuously employed is not so common, and it is even less likely for women. In fact, the proportion of individuals that during this period work more than the 90 per cent of the time is only 38 per cent for men and 22 per cent for women.

The measure of experience proposed as an alternative, **accumulated experience**, $ACCEX_{it}$, is built as the sum of a set of variables that measure the fraction of time

[10]According to data from the Encuesta de población activa (EPA), in the last trimester of 2006 the rate of temporality was 33.8 per cent. For the interval of age from 25 to 29 years it rises above 44 per cent. The number is even greater in the youngest segment (less than 25).

[11]The subsample consists on individuals that has finalized their studies and that are observed from 25 to 29 years. In total, there are 2,184 observations in the sample.

(number of months in a year) that an individual i has spent working in the last year, $X_{i(t-1)}$, two years ago, $X_{i(t-2)}$, three years ago, $X_{i(t-3)}$, ... until the beginning of her professional career, X_{i1}. That is,

$$ACCEX_{it} = X_{i(t-1)} + X_{i(t-2)} + X_{i(t-3)} + \ldots + X_{i1} = \sum_{s=1}^{t-1} X_{i(t-s)}. \tag{1.1}$$

With this measure we can easily take into account the existence of interruptions on individual labour market histories.

1.3.2 Accumulated Experience and Potential Experience

In order to establish a comparison between the two measures of experience, I estimate by OLS wage equations as proposed by Mincer (1974), separately for men and women[12]. The dependent variable, y_{it}, is the logarithm of real gross hourly wage. In addition to experience measures, I include a set of explanatory variables common in all the specifications (some have temporal variation, W_{it}, and other are constant at the individual level, Z_i): individual characteristics as birth cohort, dummy variables that indicate if the person is married or if there are children in the household, and educational level; characteristics related to the job position as tenure, type of employer, type of contract, part-time, firm size and type of occupation; and the labour market situation by means of time and region dummies[13]. Only the birth year dummies are not commonly included in empirical wage equations. I include them because there proves to be a marked decline in wages for

[12]In the empirical application, a Chow test rejected at the 5% level the null hypothesis of equality of coefficients for men and women.

[13]A detailed explanation of these variables is offered in Appendix 1.B.

successive birth cohorts. Formally,

$$y_{it} = \alpha_0 + \alpha_1 ACCEX_{it} + \alpha_2 (ACCEX_{it})^2 + \alpha_3' W_{it} + \alpha_4' Z_i + u_{it}, \tag{1.2}$$

$$y_{it} = \delta_0 + \delta_1 POTEX_{it} + \delta_2 (POTEX_{it})^2 + \delta_3' W_{it} + \delta_4' Z_i + v_{it}, \tag{1.3}$$

where the error terms, u_{it} and v_{it}, are assumed to be white noises.

As a first approximation to the relationship between the two measures of experience, if we consider a regression of the potential measure over the accumulated one, it turns out that accumulated experience explains 57 per cent of the variation in potential experience for men and 39 per cent in the case of women. If - in addition - I include the age at which individuals start working, the R^2 rises to 77 per cent for males and 67 per cent for females. These results indicate that women in the sample delay their entrance to the market and suffer more interruptions throughout their careers.

Tables A.3 and A.4 show the estimates for the specification with the accumulated experience and the potential experience measures, respectively. Looking at the two first columns of both tables (specification I), we can see that the sign of the coefficients seems the correct. As we would expect, variables related to human capital, accumulated or potential experience and - mainly - educational level, have positive and significant effect on wages[14]. Also being older or working as a civil servant, part time, in bigger firms, in positions with responsibility or as a manager or graduate, has positive effect on wages. With the potential experience measure, having higher values of tenure has positive and significant effect for both men and women.

Moreover, the presence of children at home has positive effect for males, whereas it is not significantly different from zero for females. The same happens, but with opposite sign, in the case of temporary contract. With the potential experience measure, being

[14] In the case of experience, I am referring to the joint effect of the linear and the quadratic term.

married is positively related with wages in the case of women.

Next, I consider additional variables that try to capture the importance of career interruptions (specification II, columns 3 and 4). In Table A.3, I include a variable that measures the difference between potential and accumulated experience. This difference would be positive due to two reasons: (a) if the individual does not start working just after finishing studies, and (b) if the individual interrupts her career. In order to isolate the second effect, I also include a dummy variable of late incorporation to the job market. The main result is that this difference has significant negative effect for women whereas for men it is not significantly different from zero. In a similar way, in Table A.4, I introduce a variable called interruptions that is equal to 1 if individual i at year t has worked less than 12 months (12 months is the amount assumed by the potential experience measure). Again, we can observe that the coefficient is not significant for males whereas for females has a negative and significant effect.

Another interesting feature is the comparison of wage returns to experience between the two experience variables. If these returns are different, the use of one measure or the other would have implications (specially for gender comparisons). The effect over wages of a marginal increase in experience is equal to the partial derivative with respect to experience (equations (1.2) and (1.3), respectively). In particular,

$$\alpha_2 + 2\alpha_3 * ACCEX_{it}, \tag{1.4}$$

$$\delta_2 + 2\delta_3 * POTEX_{it}. \tag{1.5}$$

Table A.5 shows the effects obtained from the previous estimates. Beginning with specification I, the main result is that with the accumulated experience measure returns to experience are higher and the gender differential decreases (at least for low levels of experience[15]). Additionally, the difference with respect to the potential experience measure is greater for women than for men. With regard to specification II, now returns to experience continue to be bigger with the accumulated measure although the distance with respect to the potential one seems to be less.

In short, using a measure of experience more accurate than the potential one has consequences. As we have seen, with the proposed measure returns to experience are larger than with the potential experience, and this is specially true for women. I use the accumulated experience measure in the rest of the work.

1.3.3 Checking Endogeneity

Since the experience variable proposed considers individual heterogeneity in the accumulation of experience, it might arise an endogeneity problem due to a correlation between this measure and unobservable wage determinants. In such a case, OLS estimates would be inconsistent. Next, I take advantage of the panel structure panel to assess this possibility.

I assume that the random error in (2), u_{it}, can be decomposed into a fixed individual component, η_i, and a random component, ϵ_{it}, both with zero mean and constant variance. Additionally, I assume - as before- that the transitory error term, ϵ_{it}, is uncorrelated with all the explanatory variables.

With regard to the individual component, η_i, a first approximation would be a fixed effects approach. Under this approach, individual heterogeneity could be arbitrarily correlated with the regressors. However, this methodology is very demanding for the sample considered here, since the time variation in the first differences of the explanatory variables is not very large. A second approximation, that represents an intermediate solution between OLS and fixed effects, consists on considering that the $\eta_i' s$ would be correlated

[15] In the sample, many of the observed labour market histories do not last more than three or four years. Due to the lack of observations, estimates for far away horizons are based mainly on extrapolations.

with some of the explanatory variables (accumulated experience and some elements of W_{it} or Z_i) but uncorrelated with the rest. This is the efficient instrumental variables method proposed by Hausman and Taylor (1981). A disadvantage of this method is that we have to impose which are the variables that are correlated with the individual effects and which are not. Only if the assumption is correct, the estimator would be consistent. Formally,

$$y_{it} = \alpha_0 + \alpha_1 ACCEX_{it} + \alpha_2 \left(ACCEX_{it}\right)^2 + \alpha'_{31} W_{it1} + \alpha'_{32} W_{it2} + \alpha'_{41} Z_{i1} + \alpha'_{42} Z_{i2} + \eta_i + \epsilon_{it}, \tag{1.6}$$

where the $\eta_i' s$ are correlated with $ACCEX_{it}$, $\left(ACCEX_{it}\right)^2$, W_{it2} and Z_{i2}, but uncorrelated with W_{it1} and Z_{i1}. The method takes the variables that are uncorrelated with the $\eta_i' s$ as instruments for the variables that do are correlated. Instruments are: (a) each endogenous variable with time variation ($ACCEX_{it}$, $\left(ACCEX_{it}\right)^2$, W_{it2}) in deviations from individual means, (b) each exogenous variable with time variation (W_{it1}) both in deviations from individual means and individual means, and (c) each exogenous variable without time variation (Z_{i1}). Therefore we get identification if we have enough exogenous variables with time variation to use as instruments for the endogenous variables that do not change.

Here, W_{it2} includes marital status, tenure, type of employer, type of contract, part time, firm size and occupation, and Z_{i2} includes education. As exogenous variables, as Booth *et al.* (2002), I consider that Z_{i1} contains birth cohort and W_{it1}, regional unemployment rate[16] and children.

Now (Table A.6), estimated coefficients for accumulated experience are slightly higher than in the previous section (more for women than for men). Those variables still have an effect significantly different from zero, like education, part time or firm size, maintain their positive relation with wages. Nevertheless, now imprecision is greater. This causes that

[16] Given the reduced regional mobility in the sample, this variable has larger time variation than region dummies. Notice that in the ECHP regional distribution is at NUTS1 level, that is not exactly the same as Autonomous Communities distribution (see variable definition in Appendix 1.B).

variables as type of contract or occupation become insignificant. Given the limited time variation in the sample as well as the fact of having to assume the scheme of correlations between the individual effects and the regressors, I take these estimations with caution[17].

1.3.4 Job Changes and Wage Growth

Job mobility is closely related to wage dynamics. I have already mentioned that the first years of the professional career concentrate a big amount of the wage growth that individuals accumulate throughout their life and mobility plays an important role on this pattern. In addition, it has been stated that among American and Italian workers this early wage growth is greater for men than for women. Time evolution for gender gross wage gap in the sample appears in Figure E.1.

Although, at labour market entry the gender wage gap is hardly perceivable, in a few years this gap become noticeable. In two years, wage growth for men is 15.94 per cent and for women 14.28 per cent. In four years, we have an accumulated growth of 26.09 per cent for males and only a 17.68 for females, whereas in six years those numbers are 44.66 and 29.28, respectively.

To analyse to what extent job changes affect wages, I include variables that indicate job changes in the wage equations[18]. Consider equation (1.3):

$$y_{it} = \delta_0 + \delta_1 POTEX_{it} + \text{second order terms and other variables} + v_{it},$$

[17] In fact, in a specification where children variable is not exogenous any more, estimates for experience coefficients are even closer to the ones obtained with OLS. However, imprecision is even greater in this case.

[18] It would be more appropriate to consider a joint model for wages and job changes since, if there exists correlation between the unobservable determinants of wages and those of job mobility, we would have a sample selection problem. A model with self-selection is out of the scope of this work and constitutes an interesting point for future research.

where

$$\begin{aligned} POTEX_{it} &= AGE_{it} - YEARS\ OF\ SCHOOLING_i - 6 \\ &= t + (AGE_{i0} - YEARS\ OF\ SCHOOLING_i - 6) = t + c_i, \end{aligned}$$

with AGE_{i0} denoting age at which individuals enter the sample and c_i an individual specific constant. If we omit second order terms and other variables (or considering y_{it} as the part of the logwage unexplained by them), we could rewrite (3) as

$$y_{it} = \delta_0 + \delta_1 t + \delta_1 c_i + v_{it},$$

and, in terms of growth rates,

$$\Delta y_{it} = \delta_1 + \Delta v_{it} \Rightarrow \hat{\delta}_1 = \overline{\Delta y}.$$

In other words, we can interpret the estimated coefficient $\hat{\delta}_1$, corresponding to the potential experience variable, as the mean wage growth. If we add interactions of potential experience and job changes, we would obtain estimates of the mean wage growth with job change.

Table A.7 shows that the mean wage growth with job change is 0.044 for men (linear term and interaction term jointly significant at 99 per cent) and only 0.008 for women (jointly significant at 90 per cent). Without job change, mean wage growth is 0.029 and 0.015, respectively. According to these estimates, the early-career wage growth of males is favoured by job mobility, but the same does not happen in the case of females.

Next section will address whether there are gender differences in the determinants of job mobility. Those differences may be causing that males and females wages do not grow at the same rate.

1.4 Job mobility

In this section, I analyse whether there are gender differences on the mobility patterns of young workers, the probabilities of each type of job change, or the factors that affect these movements. For undertaking this task, I consider all the transitions from the first job of each individual and I make distinctions with respect to the type of change (promotion, layoff and quit) and with respect to the destiny of the change (job to job or job to nonemployment, that is, unemployed or out of the labour force). I assume that an individual experiments at most one job change per year, because the reason for changing is only available for one transition each year[19].

1.4.1 Definitions

TO STAY: transition job to job, without change of employee nor duties.

PROMOTION: transition job to job, without change of employee but in a better position.

LAYOFF: transition job to job or job to nonemployment, if the reason for changing is forced by the employer, end of the fixed-term contract or by business closing.

QUIT: transition job to job or job to nonemployment, due to other reason (better position, getting married, studies, military service, illness or own inability, taking care of children or older people, ...).

[19] Notice that this is a quite restrictive assumption, since the temporality rate among the individuals in the sample is more than 40 per cent. In fact, if we counted the cases for which two or more job changes occur in a year, those changes represent around a 30 per cent of the total transitions. In any case, this is a limitation imposed by the own nature of the information available.

1.4.2 Estimation Results

I consider a multinomial logit to model the transitions across jobs and from employment to nonemployment[20]. I am interested in how *ceteris paribus* changes in the elements of a set of variables affect the probabilities of each type of change. For $j = 0, 1, 2, 3, 4, 5$, I define the probabilities $P(y = j|x)$, where now $y = \{$to stay, promotion, change job to job through layoff, change job to job through quit, change job to nonemployment through layoff, change job to nonemployment through quit$\}$, and x are personal and job characteristics that have influence on the probability of changing. The multinomial logit assumes a logistic form for those probabilities. I estimate the model by maximum likelihood[21].

In the sample there are 1470 transitions, 736 for men and 734 for women. The 55 per cent of the transitions in the case of men and the 58 per cent for women imply job changes. In fact, gender differences do not come by the side of the number of job changes they suffer. In Figure E.2 we can see that gender differences arise if we distinguish by type. Transitions job to job through layoff are higher for women and transitions job to nonemployment through quit are higher for men, although they are small in absolute terms.

Next, I consider the estimation of the multinomial logit model, separately for men and women. In a first specification I include as explanatory variables age, family (married

[20]In the empirical analysis of job mobility, discrete choice models and continuous duration models have been used. Both methodologies constitute alternative ways of modelling the same underlying process. Duration models consider the probability that a given job ends in a certain time interval conditioned on having lasted until then. Discrete choice models consider a sequence of successes or failures that are observed in each time interval, understanding by success the job change and failure, to stay in the same position. Royalty (1998) points as a main advantage of the continuous duration models the fact that the results do not depend on the considered time interval (Heckman and Singer, 1984), problem that can arise with the discrete duration and discrete choice models, in which we need to choose a given point in time when the decision takes place. Nevertheless, a model as the multinomial logit, equivalent to a discrete duration model with constant hazard rates, may have a simpler interpretation in terms of how the variables affect the probabilities of each event. For this reason, and also because many variables in the data are measure annually, I use here a discrete framework with annual intervals.

[21]For a description of the multinomial logit see Wooldridge (2001).

and/or with children), educational level, residence in regions with high economic activity, tenure and its square, accumulated experience and its square, having a position which involves supervision, working part-time and firm size[22]. In the second specification I add type of occupation[23]. Although these variables may introduce endogeneity problems, it can be of interest to take into account that men and women are concentrated in different occupations and, moreover, they can change differently due to occupational segregation.

Tables A.8-A.11 show the multinomial logit estimates. From the estimated coefficients I obtain the predicted probabilities (Table A.12 for the specification without occupations[24]), that must fit the sample proportions in Figure E.2. The interpretation of the coefficients however is not direct; that's why I also calculate marginal effects in order to know the effect that a change in a variable has on the transition probabilities. First I calculate those marginal effects at the individual level, and then I obtain the mean for the group of males and for the group of females (Tables A.13-A.18) [25].

In the probability of staying (Table A.13) the two main factors are tenure and experience. The qualitative effect is the same for males and females in both cases. As we could expect, the effect is positive for tenure. The higher the tenure the higher the probability of remaining in the same position. This result goes in line with the idea that accumulating specific human capital makes the individual more indispensable. With regard to experience, in principle it is not clear why having greater experience, for a given level of tenure and age, diminishes the probability of staying. Nevertheless, if we compared this

[22]Notice that I consider the characteristics of the original post.

[23]A detailed explanation of all these variables is included in Appendix 1.B.

[24]With regard to these predicted probabilities, conclusions are very similar in the case of the specification that includes variables of type of occupation.

[25]I comment here only the results for the specification without occupation variables, since the general conclusions do not vary. In fact, adding type of occupation becomes significantly relevant only in two cases: the probability of layoff for women moving to another job, that is lower if they work in occupations of management or related to university degrees instead of unskilled occupations (Table A.15), and the probability of quit moving to unemployment or inactivity, that is lower for men in fishing, agriculture or manufacture (Table A.18), than for men in unskilled positions.

effect with the effect that more experience has on the probability of promotion (Table A.14) or voluntary job change (Table A.16), positive in both cases and for men as well as for women, it seems that the smaller probability of staying is due to the fact that turnover increases with experience. These effects, therefore, would be compatible with the accumulation of general human capital.

For males and females the probability of promotion (Table A.14) is higher with a college degree than with only a primary school degree (this effect disappears in the specification that includes occupation). With regard to gender differences, other factors such as working part-time, working in firms of medium size or age, have a negative effect on the probability of promotion for women, whereas for men are not significant. These effects may be caused by the fact that women face non-professional restrictions that are limiting their possibilities of promoting. On the contrary, living in regions with high economic activity has a significantly positive effect on the promotion probability of women. In these areas promotion opportunities can be more numerous or, at least, more accessible.

If the job change is from a job to other through layoff (Table A.15), for both males and females, the probability of layoff is lower if they have higher tenure (it turns out costly for the employer to dismiss the worker and to hire another for replacing the first one). For men holding a position of responsibility turns out specially favourable, and for women, working in occupations that require high qualification. For males and females, greater experience is associated with a greater probability of layoff. Given the high temporality rate of the workers in the sample, a possible explanation of this result may be that accumulating experience by means of temporary contracts is not valuable for the employers[26].

In the transition job to job through quit (Table A.16), tenure and experience are important both for men and women (it was the contrary for staying). Again following a human capital perspective, the higher the tenure the lower the probability of leaving; the

[26]Layoff definition includes end of temporary contract.

higher the experience, the higher the probability of changing job.

In case of layoff to unemployment or out of the labour force, several characteristics become relevant (Table A.17). For men, having a family or living in regions with a high economic activity, reduces the probability of this transition. Any of these factors is significant for the sample of women. For females, the probability of this transition is lower when they have a higher value of tenure.

If the transition consists of leaving voluntarily a position to nonemployment, again the variables that matter for each group are different (Table A.18). For men, occupation and educational degree are important factors. They quit less if they are in a medium-skill occupation (qualified workers of agriculture or fishing) than in an unskilled occupation (labourers), or if they have a college degree with respect to those with only primary education. However, holding a position that involves responsibility increases the probability of quit. Since in the sample we have young workers who have finished their studies, some of these quits may be due to periods of inactivity dedicated to complete their professional formation. For women, quit probability is lower if they work in firms of bigger size instead of small firms (between one and four employers). It seems reasonable that in big firms labour conditions would be more flexible.

To sum up, experience, tenure and - sometimes - education, have a relevant influence over the turnover probabilities both for males and females. On the other hand, differences arise from factors, related to the job position or the social environment, that are differently important for each group. For men holding a position with responsibility, having a family or living in areas with higher economic activity turns out to be important when changing job. On the contrary, for women it is relevant the type of journey or the size of the firm.

1.5 Conclusions

This chapter analyses gender differences in labour market histories of young workers in Spain. Several theoretical hypothesis have suggested that differences in those profiles may have implications on the gender wage gap that we observe, through gender differences in the accumulation of experience and training, or through differences in job mobility.

The study focuses on two key features of individual labour market profiles: interruptions and job changes. Firstly, I propose an experience measure that - as opposed to the conventional potential experience variable - considers the existence of discontinuities in the professional career of workers. Secondly, I analyse gender differences in job mobility patterns among young workers.

From the comparison between the proposed experience measure - accumulated experience - and the one used normally - potential experience - it turns out that wage returns to experience are higher with the more accurate measure and that difference is greater for women than men. The conclusion is that it seems to exist a gender wage penalty to interruptions.

Regarding mobility, I find that turnover rates are similar for men and women. If we distinguish by type of transition, we can see that women suffer layoffs more likely than men, whereas males change job more often in a voluntary way. With respect to the variables that affect these job changes, I obtain that tenure and experience affect significantly transitions job to job and to nonemployment, but in the same direction for men and women. Differences come from the side of some characteristics that are relevant for one of the two groups and not for the other, specially in case of promotion or in the transitions to nonemployment. For men having a position with responsibility or having a family turns out to be important when changing job. On the contrary, for women it is relevant the type of journey or the size of the firm.

Finally, in addition to the gender penalty to interruptions, I also find that early-career wage growth is greater for men than for women, and this is specially true in years when job changes occur.

Appendix of Chapter 1

1.A Sample Selection

The ECHP user files are provided separately for each wave and in five different types: household files, individual files, members of the household, longitudinal connection and relationship files.

Although the analysis is based mainly on the variables included in the individual files, I need to merge them with the remain files to obtain relevant information on some family aspects. This is the case, for instance, of the variable region of residence, including in the household files, or presence of children at home, obtained as a combination of individual, household and relationship files.

After adding these variables, I append the waves. This is the unfiltered sample that represents the starting point. Then I apply successive filters until obtaining a sample in which the construction of completed individual labour markets histories is feasible. The successive steps are the following:

1. Keep only individuals aged 16 to 39 over the period.

2. Drop those with a spell of self-employment.

3. Keep only those who have finished their studies and are observable since the beginning of their professional career.

4. Drop observations previous to the first job because they are uninformative with respect to wages or labour market trajectories.

= FINAL SAMPLE (1,142 individuals and 3,251 observations).

Finally, I have young workers aged 16 to 39, who are observable since their first job, once they have finished their studies. These individuals, when they work, they do it as employees and for more than 15 hours per week. In total there are 555 males and 587 females in the sample.

1.B Definition of Variables

Real hourly wage: the ECHP includes information on gross and net average monthly income for employees. If an individual has more than one job, only the amount corresponding to the main position is included. Additionally, it provides information on the number of hours per week that the individual works in its main job. I impose a maximum of 60 hours to the number of weekly hours that an individual can work. Hourly wage is obtained as the gross monthly wage multiplied by 12 and divided by 52 to have an average weekly wage and, next, it is divided by the number of hours per week.

In order to express the hourly wage in real terms I deflate it with the Indice de Precios de Consumo (IPC) that publishes the INE. I use the series of annual means of the general index with base 1992. An additional correction consists of eliminating the observations corresponding to individuals whose hourly wage is below the minimum wage[27].

Potential Experience: it is defined as age minus years of schooling minus 6, that is, current age minus age when individuals finished studies.

Accumulated Experience: it is constructed as the sum of a set of variables, $X's$, that measure, for a given individual i and time t, the fraction of time (number of months in a year) that this individual has spent working in the last year, $X_{i(t-1)}$, two years ago, $X_{i(t-2)}$, three years ago, $X_{i(t-3)}$, ... until the beginning of her professional career, X_{i1}.

[27]This correction affects 3.7 per cent of men and 5.2 percent of women.

Educational level: dummies defined for the highest degree obtained by the individual (primary education, graduate, college).

Tenure: it is constructed from the answers that individuals give when they are asked in which year they began to work with the present employer. It is obtain as the difference between the current year and the year the individual begins to work with the present employer. It is used as a continuous variable, or as dummies of less than 1 year, from 1 to 2 years, and more than 2 years of tenure.

Personal characteristics: age (continuous variable), sex (two dummies), marital status (married as opposed to another situation), presence of children at home (dummy variable), family (married and/or with children).

Economic Centres: in the ECHP regional division is at NUTS1 level.

- Northwest: Galicia, Asturias and Cantabria.

- Northeast: Basque Country, Navarra, Rioja, and Aragon.

- Madrid.

- Centre: Castilla - Leon, Castilla la Mancha, Extremadura.

- East: Catalonia, Valencia, The Balearics Islands.

- South: Andalusia, Murcia, Ceuta and Melilla.

- The Canary Islands.

Centres is a dummy variable that takes the value 1 if an individual lives in Madrid, Catalonia or Basque Country, regions with a higher economic activity.

Type of contract: temporary or permanent work.

Type of journey: part-time or full-time work.

Degree of responsibility: dummy variable whether a position involves supervision duties.

Firm size: from 1 to 4 employees, from 5 to 49 employees, and 50 or more employees.

Occupation: I have grouped the variable occupation in four categories. Since the results can be sensible to this grouping, the groups are establish based on similar requirements on qualification and responsibility. The four categories are:

MANAGERS and PROFESSIONALS: Directors of the Public Administrations, Professions associated to college degrees in the fields of pure and natural sciences, health and education, Professionals of the Law, Social sciences and humanities, Technical experts on pure and natural sciences, health and education and Professionals of support in financial, commercial operations and in the administrative management.

CLERICAL and SERVICES: Clerical employees and workers of catering and personal services, protection and security, and sales workers

AGRICULTURE and MANUFACTURE: Qualified workers in agriculture and fishing, qualified workers on construction, extractive industries, food, drinks and tobacco, wood and textile industry, qualified craftsmen and workers in the metallurgy, operators and fitters of industrial machinery, and transport.

UNSKILLED: Non-qualified services and commerce workers, farming and fishing labourers, labourers of mining industry, manufacturing construction, industries and transport.

Transitions: categorical variable that takes six values, one for each transition (according to the definitions included in section 1.4.1): staying, promotion, change job to job through layoff, change job to nonemployment through layoff, change job to job through quit, and change job to nonemployment through quit.

Time effects: eight dummies, one for each year.

Chapter 2

Modelling Heterogeneity and Dynamics in the Volatility of Individual Wages

2.1 Introduction

Estimates of individual earnings processes are useful for a variety of purposes, which include testing between different models of the determinants of earnings distributions, building predictive earnings distributions, or calibrating consumption and saving models.

While several papers have focused on modelling the heterogeneity and time series properties of the conditional mean of earnings given its past (Lillard and Willis, 1978; MaCurdy, 1982; Abowd and Card, 1982, among others), the modelling of the conditional variance has been mostly neglected. However, in many applications it is important to understand the behavior of higher order moments of the process. This would be the case if we consider an individual trying to forecast her future earnings, in order to guide savings or other decisions. As the individual faces various sorts of uncertainty, we shall be interested in forecasting not only the level of earnings but also its variance. The properties of the variance will be important for describing wage profiles over time and for better understanding what drives fluctuations in them. A richer specification can contribute also to modelling choices in models that use the earnings process as an input. In fact, recent studies stress the relevance of considering a variance that varies with time and across individuals (Meghir and Windmeijer, 1999; Chamberlain and Hirano, 1999; Meghir and Pistaferri, 2004; Albarrán, 2004; Alvarez and Arellano, 2004).

There are also many papers that study the increase in the cross-sectional variance of earnings since the 70's until today (Juhn, Murphy, and Pierce, 1993, and many others). This growth in the aggregate variance is associated with an increase in inequality. Much less is known about the behaviour of the conditional variance given observed and unobserved individual characteristics.

In this chapter, I propose a likelihood-based panel data model for the heterogeneity and dynamics of the conditional mean and the conditional variance of individual wages. In particular, I build a dynamic panel data model with linear individual effects in the mean and multiplicative individual effects in the conditional ARCH type variance function. Therefore, with this model, we can say to what extent the time evolution of the variance is determined by permanent individual heterogeneity or by state dependence effects. This distinction would be crucial, for instance, in the case of precautionary savings as the consumer would behave differently if she knows that the risk she suffers is permanently higher, than if it is only due to a period of higher volatility.

It is well known that failure to control for individual unobserved heterogeneity can lead to misleading conclusions. This problem is particularly severe when the unobserved heterogeneity is correlated with explanatory variables. Such a situation arises naturally in a dynamic context. Here, I adopt a fixed effects perspective leaving the distribution for the unobserved heterogeneity completely unrestricted and treating each effect as one different parameter to be estimated.

There is an extensive literature on how to estimate linear panel data models with fixed effects (see Chamberlain, 1984, and Arellano and Honoré, 2001, for references), but there are no general solutions for non-linear cases. If the number of individuals N goes to infinity while the number of time periods T is held fixed, estimation of non-linear models with fixed effects by maximum likelihood suffers from the so-called Incidental Parameters Problem (Neyman and Scott, 1948). This problem arises because the unobserved individual characteristics are replaced by inconsistent sample estimates, which biases estimates of model parameters. In particular, the bias of the maximum likelihood estimator is of order $1/T$. The number of periods available for many panel data sets is such that it is not less natural to talk of time-series finite sample bias than of fixed-T inconsistency or underidentification. In this light, an alternative reaction to the fact that micro panels are short is to ask for approximately unbiased estimators as opposed to estimators with no bias at all. This approach has the potential of overcoming some of the fixed-T identification difficulties and the advantage of generality. Methods of estimation of nonlinear fixed effects panel data models with reduced bias properties have been recently developed (see Arellano and Hahn, 2006a, for a review). There are automatic methods based on simulation (Hahn and Newey, 2004), bias correction based on orthogonalization (Cox and Reid, 1987; Lancaster, 2002) and their extensions (Woutersen, 2002; Arellano, 2003), analytical bias correction of estimators (Hahn and Newey, 2004; Hahn and Kuersteiner, 2004), bias correction of the moment equation (Carro, 2006; Fernández-Val, 2005) and bias corrections for the concentrated likelihood (DiCiccio and Stern, 1993; Severini, 1998a; Pace and Salvan, 2005).

Following this perspective, I build a modified likelihood function for estimation and inference. Using a bias-corrected concentrated likelihood makes it possible to reduce the estimation bias to a term of order $1/T^2$, without increasing its asymptotic variance. This is very encouraging since the goal is not necessarily to find a consistent estimator for fixed T, but one with a good finite sample performance and a reasonable asymptotic approximation for the samples used in empirical studies.

The contributions of the chapter are twofold. First, I develop several versions of the modified likelihood based on DiCiccio and Stern (1993), Severini (1998a), Pace and Salvan (2005), and Arellano and Hahn (2006b) adapted to a dynamic conditional variance model. Second, I show how this approach works in practice for a specific empirical setting. The small sample performance of bias corrected estimators is investigated in a Monte Carlo study. The simulation results show that the bias of the maximum likelihood estimator is substantially corrected for samples designs that are broadly calibrated to the one used in the empirical application. The empirical analysis is conducted on data drawn from the 1968-1993 Panel Study of Income Dynamics (PSID). These models and data are interesting because we do not know much how the volatilities of individual wages behave in a period of increasing aggregate inequality. I find that it is important to account for individual unobserved heterogeneity and dynamics in the variance, and that the latter is driven by job mobility. I also find that the model explains the non-normality observed in logwage data.

In a similar sample for male earnings, Meghir and Pistafferi (2004) find strong evidence of state dependence effects as well as evidence of unobserved heterogeneity in the variances[1]. They also propose an autoregressive conditional heteroskedasticity panel data model of earnings dynamics, but they separate into a permanent component and a transitory component of earnings shocks. This can be appropriate in models where the author makes assumptions about the nature of the different shocks that affect the income process. Nevertheless, a model with a permanent component $I(1)$ imposes a unit root, i.e., a value for the autoregressive coefficient in the mean equal to one, whereas recent evidence suggests a value for this coefficient around $0.4 - 0.5$ (Alvarez and Arellano, 2004). I use

[1] Also Lin (2005), using a subsample of the dataset considered by Meghir and Pistaferri (2004), finds statistically significant evidence of ARCH effects in earnings dynamics. He considers an ARCH-fixed effects estimator in a “quasi-lineal” setting. Here we consider a different econometric framework, which let us handle models with multiple effects and estimators without being constrained to the availability of differencing schemes.

a single-shock, multiple effects model instead[2]. This parsimonious specification would be useful for describing and estimating wage distributions (Chamberlain and Hirano, 1999). Meghir and Pistaferri recover orthogonality conditions for the estimation. Their method depends critically on the linear specification for the variance. But even in this case, they recognize that they cannot do fixed-T consistent GMM estimation because they have weak instruments. So, they implement a WG-GMM estimator which is only consistent when $T \to \infty$. What is specially worried about this is that they have a bias of order $1/T$ as opposed to my estimator which has a bias of order $1/T^2$. This difference is very important, as we will see in the simulations with respect to the MLE which also has a bias of order $1/T$. Even worse, because the WG-GMM estimator use arbitrary moment conditions and thus it is less efficient than MLE. I choose an exponential specification that implies a conditional variance always nonnegative regardless of the parameter values and in addition it has a steady-state distribution (Nelson, 1992). What is interesting is that the estimation method does not depend on the particular specification. It could also use without major changes a quadratic specification as the one of Meghir and Pistaferri.

Two limitations of the model are the following: (i) so far there is not adjustment for measurement error; and (ii) there is not explicit treatment of job changes. It is known that measurement error is important for PSID wages and that part of the wages variance may be due to job mobility, so these issues need to be addressed in further work.

The rest of the chapter is organised as follows. Section 2.2 presents the model and the likelihood function. Section 2.3 reviews the alternative approaches for correcting the likelihood adapted to this particular setting. Section 2.4 shows some simulations to study the finite sample performance of the bias corrections for the concentrated likelihood. In Section 2.5, I present the empirical application on individual wages and in Section 2.6 the implications of the model for consumption growth. Section 2.7 concludes.

[2] Meghir and Windmeijer (1999) and Albarrán (2004) use single-shock models as well but they do not have an application to data.

2.2 The Model and the Likelihood Function

2.2.1 The Model

I consider the following model of standardized logwages where i and t index individuals and time, respectively:[3]

$$y_{it} = \alpha y_{it-1} + \eta_i + e_{it} = \alpha y_{it-1} + \eta_i + h_{it}^{1/2}\epsilon_{it};\ \ (i = 1, ..., N; t = 1, ..., T)$$

with

$$E\left(y_{it}|y_i^{t-1}, \Theta_i\right) = \alpha y_{it-1} + \eta_i,$$

and

$$\begin{aligned} h_{it} &= Var\left(y_{it}|y_i^{t-1}, \Theta_i\right) = E\left(e_{it}^2|y_i^{t-1}, \Theta_i\right) \\ &= \exp\left(\psi_i + \beta\left[|\epsilon_{it-1}| - E\left(|\epsilon_{it-1}|\right)\right]\right) \\ &= h\left(\epsilon_{it-1}, \psi_i\right). \end{aligned}$$

In these expressions, $\{y_{i0}, ..., y_{iT}\}_{i=1}^N$ are the observed data, $\Theta_i = (\eta_i, \psi_i)'$ are the individual unobserved fixed effects, e_{it} is an ARCH process, and $\{\epsilon_{it}\}$ is an *i.i.d.* sequence with zero mean and unit variance[4]. The log formulation implies that h_{it} is always nonnegative, regardless of the parameter values (Nelson, 1992). Finally, I denote the vector of common parameters as $\Gamma = (\alpha, \beta)'$.

For the conditional mean, I consider an autoregressive specification where the parameter α measures the persistence on the level of wages to shocks, η_i describe permanent

[3]In the sequel, for any random variable (or vector of variables) Z, z_{it} denotes observation for individual i at period t, and $z_i^t = \{z_{i0}, ..., z_{it}\}$, i.e. the set of observations for individual i from the first period to period t.

[4]In the empirical analysis, I approximate the absolute value function by means of a differentiable function.

unobserved heterogeneity and e_{it} reflects shocks that individuals receive every period. Departing for the classical AR(1) process, I permit that the variances, given past observations, change over time and across individuals. This particular ARCH type specification allows me to capture two patterns of wage volatility. The first one is individual heterogeneity, ψ_i: wage volatilities of different individuals can vary differently. For instance, there can be different variances of wages between civil servants and workers of a sales department and also between workers of sales departments in big and small firms. The second one is dynamics, β, reflecting that periods of high volatility in wages tend to be consecutive and vice versa. This feature would be noticeable not only for sellers, but also for funds managers or, in general, for workers that receive bonuses.

2.2.2 The Likelihood Function

Under the assumption that $\epsilon_{it} \sim N(0,1)$, that is, $\epsilon_{it}|y_i^{t-1}, \Theta_i \sim N(0,1)$ then, conditional on the past, the model is normal heteroscedastic

$$y_{it}|y_i^{t-1}, \Theta_i \sim N(\alpha y_{it-1} + \eta_i, h_{it}),$$

and the individual likelihood, conditioned on initial observations, and fixed effects, is

$$f\left(y_{i1}, ..., y_{iT}|y_{i0}, \Theta_{i0}\right) = \prod_{t=1}^{T} f\left(y_{it}|y_{it-1}, \Theta_{i0}, \Gamma_0\right).$$

The log-likelihood for one observation, ℓ_{it}, differs from the linear model with normal errors through the time-dependence of the conditional variance. For any individual i and $t > 1$, we can write

$$\ln f\left(y_{it}|y_{it-1}, \Theta_i, \Gamma\right) = \ell_{it}\left(\Gamma, \Theta_i\right) \propto -\frac{1}{2}\ln\left(h\left(\epsilon_{it-1}, \psi_i\right)\right) - \frac{1}{2}\frac{\left(y_{it} - \alpha y_{it-1} - \eta_i\right)^2}{h\left(\epsilon_{it-1}, \psi_i\right)}.$$

Initial conditions. Evaluation of the likelihood at $t = 1$ requires pre-sample values for ϵ_{it}^2 and h_{it}. For $t = 1$,

$$y_{i1} = \alpha y_{i0} + \eta_i + [h(\epsilon_{i0}, \psi_i)]^{1/2} \epsilon_{i1},$$

where $h(\epsilon_{i0}, \psi_i) = h(y_{i0}, y_{i,-1}, y_{i,-2}, ...)$. This is a model for $f(y_{i1}|y_{i0}, y_{i,-1}, y_{i,-2}, ..., \Theta_{i0})$ or for $f(y_{i1}|y_{i0}, \epsilon_{i0}, \Theta_{i0})$ where ϵ_{i0} resumes all the past values of y_{it}, but what we would need is $f(y_{i1}|y_{i0}, \Theta_{i0})$. Since,

$$E(y_{i1}|y_{i0}, \Theta_{i0}) = E(y_{i1}|y_{i0}, \epsilon_{i0}, \Theta_{i0}) = \alpha y_{i0} + \eta_i,$$

and

$$\begin{aligned}
Var(y_{i1}|y_{i0}, \Theta_{i0}) &= E(h(\epsilon_{i0}, \psi_i)|y_{i0}, \Theta_{i0}) + Var(\alpha y_{i0} + \eta_i|y_{i0}, \Theta_{i0}) \\
&= E(h(\epsilon_{i0}, \psi_i)|y_{i0}, \Theta_{i0}) + Var(\eta_i|y_{i0}, \Theta_{i0}) \\
&= \varphi(\eta_i, \psi_i, \Gamma).
\end{aligned}$$

Thus, $f(y_{i1}|y_{i0}, \Theta_{i0})$ would be a mixture given that:

$$f(y_{i1}|y_{i0}, \Theta_{i0}) = \int f(y_{i1}|y_{i0}, \epsilon_{i0}, \Theta_{i0})\, dG(\epsilon_{i0}|y_{i0}, \Theta_{i0}).$$

For simplicity, I consider an approximate model where $y_{i1}|y_{i0}, \Theta_{i0} \sim N(\alpha y_{i0} + \eta_i, h_{i1})$ and, as suggested by Bollerslev (1986), I use the mean of the squared residuals as an estimate for $h_{i1} = \frac{1}{T}\sum_{t=1}^{T} e_{it}^2$.[5] As $T \to \infty$, h_{i1} is the steady-state unconditional variance of e_{it} given fixed effects, that is,

$$\varphi(\eta_i, \psi_i, \Gamma) = p\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} (y_{it} - \alpha y_{it-1} - \eta_i)^2.$$

[5] Another alternative would be adding the missing variances as parameters to be estimated.

Let the individual likelihood function be

$$\pounds_i\left(\Gamma,\Theta_i\right)=\prod_{t=2}^{T}\frac{1}{\left[h\left(\epsilon_{it-1},\psi_i\right)\right]^{1/2}}\phi\left(\frac{y_{it}-\alpha y_{it-1}-\eta_i}{\left[h\left(\epsilon_{it-1},\psi_i\right)\right]^{1/2}}\right)\cdot\frac{1}{\left[h_{i1}\right]^{1/2}}\phi\left(\frac{y_{i1}-\alpha y_{i0}-\eta_i}{\left[h_{i1}\right]^{1/2}}\right),$$

and the log-likelihood of each observation

$$\ell_{it}\left(\Gamma,\Theta_i\right)=-\frac{1}{2}\ln\left(h_{it}\right)-\frac{1}{2}\frac{\left(y_{it}-\alpha y_{it-1}-\eta_i\right)^2}{h_{it}},$$

where

$$h_{it}=\begin{cases}\frac{1}{T}\sum\limits_{t=1}^{T}e_{it}^2 & if\quad t=1,\\ \exp\left(\psi_i+\beta\left[\left|\epsilon_{it-1}\right|-E\left(\left|\epsilon_{it-1}\right|\right)\right]\right) & if\quad t>1.\end{cases}$$

2.3 Correcting the Likelihood Function

In this section, I adopt a likelihood-based approach that allows me to deal with dynamics and multiple fixed effects in the estimation. The MLE of Γ, concentrating out the Θ_i, is the solution to

$$\widehat{\Gamma}\equiv\arg\max_{\Gamma}\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\ell_{it}\left(\Gamma,\widehat{\Theta}_i\left(\Gamma\right)\right);\ \widehat{\Theta}_i\left(\Gamma\right)\equiv\arg\max_{\Theta}\frac{1}{T}\sum_{t=1}^{T}\ell_{it}\left(\Gamma,\Theta\right).$$

Incidental Parameters Problem. In this context, fixed effects MLE suffers from the incidental parameters problem noted by Neyman and Scott (1948). In this case, the incidental parameters would be the individual effects. The problem arises because the unobserved individual effects Θ_i are replaced by sample estimates $\widehat{\Theta}_i\left(\Gamma\right)$: as only a finite number T of observations are available to estimate each Θ_i, the estimation error of $\widehat{\Theta}_i\left(\Gamma\right)$ does not vanish as the sample size N grows, and this error contaminates the estimates of common parameters in nonlinear models. Let

$$L(\Gamma) \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} E\left[\sum_{t=1}^{T} \ell_{it}\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right)\right].$$

Then, from the usual maximum likelihood properties, for $N \to \infty$ with T fixed, $\widehat{\Gamma}_T = \Gamma_T + o_p(1)$, where $\Gamma_T \equiv \arg\max_{\Gamma} L(\Gamma)$. In general, $\Gamma_T \neq \Gamma_0$, but $\Gamma_T \to \Gamma_0$ as $T \to \infty$.

Due to the noise in estimating $\widehat{\Theta}_i(\Gamma)$, the expectation of the concentrated likelihood is not maximized at the true value of the parameter. This problem can be avoided by correcting the concentrated likelihood.

The bias in the expected concentrated likelihood at an arbitrary Γ can be expanded in orders of magnitude of T

$$E\left[\frac{1}{T}\sum_{t=1}^{T} \ell_{it}\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right) - \frac{1}{T}\sum_{t=1}^{T} \ell_{it}\left(\Gamma, \bar{\Theta}_i(\Gamma)\right)\right] = \frac{b_i(\Gamma)}{T} + o\left(\frac{1}{T}\right),$$

where $\bar{\Theta}_i(\Gamma)$ maximizes $\lim_{T\to\infty} E\left[T^{-1}\sum_{t=1}^{T} \ell_{it}(\Gamma, \Theta)\right]$. As it is shown in Appendix 2.A, the form of the approximate bias of the concentrated likelihood is:

$$\frac{b_i(\Gamma)}{T} \approx \frac{1}{2} tr\left(H_i(\Gamma) Var\left[\widehat{\Theta}_i(\Gamma)\right]\right) = \frac{1}{2T} tr\left(H_i^{-1}(\Gamma)\Upsilon_i(\Gamma)\right),$$

where

$$\begin{aligned} H_i(\Gamma) &= -\lim_{T\to\infty} E\left[\frac{\partial^2 \ell_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)}{\partial\Theta_i \partial\Theta_i'}\right], \\ \Upsilon_i(\Gamma) &= \lim_{T\to\infty} TE\left[\frac{\partial \ell_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)}{\partial\Theta_i} \cdot \frac{\partial \ell_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)}{\partial\Theta_i'}\right], \end{aligned}$$

and

$$\ell_i(\Gamma, \Theta_i) = \frac{1}{T}\sum_{t=1}^{T} \ell_{it}(\Gamma, \Theta).$$

For further discussion on the estimation method and a formal analysis of the asymptotic properties of the bias-corrected estimators when N and T grow at the same rate see Arellano and Hahn (2006b).

In this chapter, I consider three alternative estimators of Γ which maximize a bias-corrected concentrated likelihood function like:

$$\begin{aligned}\widetilde{\Gamma} &= \arg\max_{\Gamma} \frac{1}{N}\sum_{i=1}^{N} \ell_{mi}\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right) \\ &= \arg\max_{\Gamma} \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{T}\sum_{t=1}^{T} \ell_{it}\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right) - \frac{1}{T}\hat{b}_i(\Gamma)\right].\end{aligned}$$

Letting $\hat{b}_i(\Gamma)$ be an estimated bias, $\widetilde{\Gamma}$ is expected to be less biased than the MLE $\widehat{\Gamma}$. Moreover, in a likelihood context, it is appropriate to consider a local version of the estimated bias using that at the truth $H_i^{-1}(\Gamma_0)\Upsilon_i(\Gamma_0) = 1$ (Pace and Salvan, 2005). As it is shown at the end of Appendix 2.A, this local version of $\widehat{b}_i(\Gamma)$ gives

$$\widehat{b}_i(\Gamma) = -\frac{1}{2}\ln\det \widehat{H}_i(\Gamma) + \frac{1}{2}\ln\det \widehat{\Upsilon}_i(\Gamma).$$

In practice, for estimating the bias I need to estimate the hessian term, $H_i(\Gamma)$, and the expected outer product term, $\Upsilon_i(\Gamma)$. For estimating the first one I use its sample counterpart:

$$\widehat{H}_i(\Gamma) = -\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ell_{it}\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right)}{\partial\Theta_i\partial\Theta_i'}.$$

With regard to $\Upsilon_i(\Gamma)$, note that since

$$\frac{1}{T}\sum\nolimits_{t=1}^{T}\frac{\partial \ell_{it}\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right)}{\partial\Theta_i} = 0,$$

also

$$\frac{1}{T}\sum\nolimits_{t=1}^{T}\sum\nolimits_{s=1}^{T}\frac{\partial \ell_{it}\left(\Gamma,\widehat{\Theta}_i\left(\Gamma\right)\right)}{\partial \Theta_i}\cdot\frac{\partial \ell_{is}\left(\Gamma,\widehat{\Theta}_i\left(\Gamma\right)\right)}{\partial \Theta_i'}=0,$$

so that using the observed quantities evaluated at $\widehat{\Theta}_i\left(\Gamma\right)$ will not work. The three different corrections, presented below, are based on three different estimators for this second term of the bias.

2.3.1 Determinant Based Approach Using Expected Quantities

This approach is based on the expectation

$$\begin{aligned}&\bar{\Upsilon}_i\left(\Gamma,\Theta_i;\Gamma_0,\Theta_{i0}\right)\\ \equiv\ & TE_{\{\Gamma_0,\Theta_{i0}\}}\left[\left[\frac{\partial \ell_i\left(\Gamma,\Theta_i\right)}{\partial\Theta_i}-E\left(\frac{\partial \ell_i\left(\Gamma,\Theta_i\right)}{\partial\Theta_i}\right)\right]\cdot\left[\frac{\partial \ell_i\left(\Gamma,\Theta_i\right)}{\partial\Theta_i'}-E\left(\frac{\partial \ell_i\left(\Gamma,\Theta_i\right)}{\partial\Theta_i'}\right)\right]\right]\end{aligned}$$

obtained using the true density $f\left(y_i|y_{i0},\Gamma_0,\Theta_{i0}\right)$. Notice that in this case (for an arbitrary (Γ,Θ_i)), centering the expected outer product term is crucial because only for $E\left(\frac{\partial \ell_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)}{\partial\Theta_i}\right)$ this expectation is equal to zero. Also it is important to note that this expected quantity can be obtained for given values of (Γ,Θ_i) and (Γ_0,Θ_{i0}), analytically or numerically, because in the likelihood context the density of the data is available. However, it cannot be calculated at (Γ_0,Θ_{i0}) because true values are unknown. The estimator solves this problem replacing (Γ_0,Θ_{i0}) by their ML estimates $\left(\widehat{\Gamma},\widehat{\Theta}_i\right)$. This give us the useful quantity: $\bar{\Upsilon}_i\left(\Gamma,\widehat{\Theta}_i\left(\Gamma\right);\widehat{\Gamma},\widehat{\Theta}_i\right)$. It can be regarded as a dynamic version of Severini (1998a) or DiCiccio and Stern (1993) approximations to the modified profile likelihood.

Iterated Bias-Corrected Likelihood Estimation. An undesirable feature of this approach is its dependence on $\widehat{\Gamma}$, which may have a large bias. This problem can be avoided by considering an iterative procedure. That is, once a first corrected estimate is available,

$$\widetilde{\Gamma}_I = \arg\max_{\Gamma} \frac{1}{N} \sum_{i=1}^{N} \ell_{mi} \left(\Gamma, \widehat{\Theta}_i (\Gamma) ; \widehat{\Gamma}, \widehat{\Theta}_i \right),$$

I could use it to calculate a second one:

$$\widetilde{\Gamma}_{II} = \arg\max_{\Gamma} \frac{1}{N} \sum_{i=1}^{N} \ell_{mi} \left(\Gamma, \widehat{\Theta}_i (\Gamma) ; \widetilde{\Gamma}_I, \widehat{\Theta}_i \left(\widetilde{\Gamma}_I \right) \right).$$

Pursuing the iteration,

$$\widetilde{\Gamma}_K = \arg\max_{\Gamma} \frac{1}{N} \sum_{i=1}^{N} \ell_{mi} \left(\Gamma, \widehat{\Theta}_i (\Gamma) ; \widetilde{\Gamma}_{K-1}, \widehat{\Theta}_i \left(\widetilde{\Gamma}_{K-1} \right) \right),$$

until convergence, it would be possible to obtain an estimator $\widetilde{\Gamma}_\infty$ that solves

$$\sum_{i=1}^{N} q_{mi} \left(\Gamma, \widehat{\Theta}_i (\Gamma) ; \Gamma, \widehat{\Theta}_i (\Gamma) \right) = 0,$$

where $q_{mi} (\Gamma, \Theta_i; \Gamma_0, \Theta_{i0})$ denotes the score of $\ell_{mi} (\Gamma, \Theta_i; \Gamma_0, \Theta_{i0})$ for fixed Γ_0 and Θ_{i0}.

2.3.2 Determinant Based Approach Using Bootstrap

The first step consists in generating parametric bootstrap samples $\{y_{i1}^m, ..., y_{iT}^m\}_{i=1}^N$ with $(m = 1, ..., M)$ from the model $\left\{ \prod_{t=1}^T f \left(y_{it} | y_{i0}, \widehat{\Gamma}, \widehat{\Theta}_i \right) \right\}_{i=1}^N$ and, then, calculating the corresponding fixed effects estimates $\left\{ \widehat{\Theta}_i^m (\Gamma) \right\}_{m=1}^M$. This approach, close to Pace and Salvan (2005), is based on using a bootstrap estimate of $Var \left[\widehat{\Theta}_i (\Gamma) \right]$ given by

$$\widehat{Var} \left[\widehat{\Theta}_i (\Gamma) \right] = \frac{1}{M} \sum_{m=1}^{M} \left[\widehat{\Theta}_i^m (\Gamma) - \widehat{\Theta}_i (\Gamma) \right]^2.$$

2.3.3 Trace Based Approach for Pseudo Likelihoods

Since $\Upsilon_i\left(\Gamma,\widehat{\Theta}_i(\Gamma)\right)=0$, a trimmed version of $\Upsilon_i(\Gamma)$ might work. That is,

$$\widehat{\Upsilon}_i(\Gamma)=\Omega_0+\sum_{l=1}^{r}\left(\Omega_l+\Omega_l'\right),$$

$$\Omega_l=\frac{1}{T-l}\sum_{t=l+1}^{T}\left(1-\frac{l}{r+1}\right)\frac{\partial\ell_{it}\left(\Gamma,\widehat{\Theta}_i(\Gamma)\right)}{\partial\Theta}\cdot\frac{\partial\ell_{it-l}\left(\Gamma,\widehat{\Theta}_i(\Gamma)\right)}{\partial\Theta'}.$$

In principle r could be chosen as a suitable function of T to ensure bias reduction but, given that in practice T will be small and that the procedure is known to fail for values of r at both ends of the admissible range ($r=0$ and $r=T-1$), in practice r will be chosen equal to 2 or 3.

2.4 Monte Carlo Evidence

The practical importance of these bias corrections depends on how much bias is removed for the relatively small T that is often relevant in econometric applications.

In this section, I provide some simple versions of the model showing that these corrections can remove a large part of the bias even with small T.

2.4.1 The linear dynamic panel model with fixed effects

Consistent estimates of α for fixed T are available in the AR(1) case. I consider this model first to compare the bias correcting estimators described above with the one proposed by Lancaster (2002).

The model design is

$$\begin{aligned} y_{it} &= \alpha y_{it-1} + \eta_i + \epsilon_{it}, \quad (t = 1, ..., T;\ i = 1, ..., N) \\ \epsilon_{it} &\sim N(0,1), \quad \eta_i \sim N(0,1), \\ y_{i0} &\sim N\left(\frac{\eta_i}{(1-\alpha)}, \frac{1}{(1-\alpha^2)}\right). \end{aligned}$$

The data are generated for $T = 8$ and 16, $N = 500$ and 1000, and for $\alpha = 0.5$, and 0.8. I have simulated samples for different samples sizes because I expect the modified MLE to improve much more with T than with N. And I have also simulated samples for different values of α because the larger the α the greater the serial correlation of y_{it}, thus I expect that the estimator performs worse.

Here the MLE of α is

$$\hat{\alpha} \equiv \arg\max_{\alpha} \frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{T} \sum_{t=1}^{T} \ell_{it}\left(\alpha, \hat{\eta}_i\left(\alpha\right)\right) \right] = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{y}_{it} \tilde{y}_{it-1}}{\sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{y}_{it-1}^2},$$

where

$$\hat{\eta}_i\left(\alpha\right) \equiv \arg\max_{\eta} \frac{1}{T} \sum_{t=1}^{T} \ell_{it}\left(\alpha, \eta\right) = \bar{y}_i - \alpha \bar{y}_{i(-1)},$$

and $\bar{y}_i = \frac{1}{T} \sum\limits_{t=1}^{T} y_{it}$, $\bar{y}_{i(-1)} = \frac{1}{T} \sum\limits_{t=1}^{T} y_{it-1}$, $\tilde{y}_{it} = y_{it} - \bar{y}_i$, $\tilde{y}_{it-1} = y_{it-1} - \bar{y}_{i(-1)}$. I also consider several bias-correcting estimators of α that are obtained by maximization of a modified concentrated log likelihood like

$$\widetilde{\alpha} \equiv \arg\max_{\alpha} \frac{1}{N} \sum_{i=1}^{N} \ell_{mi}\left(\alpha, \widehat{\eta}_i\left(\alpha\right)\right).$$

- Determinant Based Approach Using Expected Quantities: in this case,

$$\widehat{H}_i\left(\alpha\right) = -\frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \ell_{it}\left(\alpha, \widehat{\eta}_i\left(\alpha\right)\right)}{\partial \eta^2} = 1,$$

$$
\begin{aligned}
\bar{\Upsilon}_i\left(\alpha,\eta_i;\alpha_0,\eta_{i0}\right) &= TE_0\left[\left.\left[\frac{\partial \ell_i\left(\alpha,\eta\right)}{\partial \eta_i}-E\left(\frac{\partial \ell_i\left(\alpha,\eta\right)}{\partial \eta_i}\right)\right]^2\right| y_{i0}\right] \\
&= TVar_0\left[\left.\frac{\partial \ell_i\left(\alpha,\eta\right)}{\partial \eta_i}\right| y_{i0}\right] = TVar\left[\bar{v}_i | y_{i0}\right].
\end{aligned}
$$

where $\bar{v}_i = \frac{1}{T}\sum\limits_{t=1}^{T}\frac{\partial \ell_{it}(\alpha,\eta)}{\partial \eta}$, [6] and as it is shown in Appendix 2.B

$$
\bar{\Upsilon}_i\left(\alpha,\eta;\alpha_0,\eta_0\right) = 1+T\left(\alpha_0-\alpha\right)^2\omega_T\left(\alpha_0\right)+2T\left(\alpha_0-\alpha\right)\psi_T\left(\alpha_0\right),
$$

with

$$
\begin{aligned}
\omega_T\left(\alpha_0\right) &= \frac{1}{T^2}\left[1+\left(1+\alpha_0\right)^2+\left(1+\alpha_0+\alpha_0^2\right)^2+\ldots+\left(1+\alpha_0+\ldots+\alpha_0^{T-2}\right)^2\right], \\
\psi_T\left(\alpha_0\right) &= \frac{1}{T^2}\left[\left(1+\alpha_0+\ldots+\alpha_0^{T-2}\right)+\left(1+\alpha_0+\ldots+\alpha_0^{T-3}\right)+\ldots+1\right].
\end{aligned}
$$

Thus

$$
\bar{\Upsilon}_i\left(\alpha,\widehat{\eta}_i\left(\alpha\right);\hat{\alpha},\hat{\eta}_i\right) = 1+T\left(\hat{\alpha}-\alpha\right)^2\omega_T\left(\hat{\alpha}\right)+2T\left(\hat{\alpha}-\alpha\right)\psi_T\left(\hat{\alpha}\right).
$$

It follows that in this case

$$
\ell_{mi}\left(\alpha,\hat{\eta}_i\left(\alpha\right);\hat{\alpha},\hat{\eta}_i\right) = -\frac{1}{2T}\sum_{t=1}^{T}\left(y_{it}-\alpha y_{it-1}-\hat{\eta}_i\left(\alpha\right)\right)^2-\frac{1}{2T}\ln\bar{\Upsilon}_i\left(\alpha,\hat{\eta}_i\left(\alpha\right);\hat{\alpha},\hat{\eta}_i\right).
$$

- Determinant Based Approach Using a Parametric Bootstrap Estimate of $Var\left[\hat{\eta}_i\left(\alpha\right)\right]$: now

$$
\ell_{mi}\left(\alpha,\hat{\eta}_i\left(\alpha\right)\right) = -\frac{1}{2T}\sum_{t=1}^{T}\left(y_{it}-\alpha y_{it-1}-\hat{\eta}_i\left(\alpha\right)\right)^2-\frac{1}{2}\ln\widehat{Var}\left[\hat{\eta}_i\left(\alpha\right)\right],
$$

where

$$
\widehat{Var}\left[\hat{\eta}_i\left(\alpha\right)\right] = \frac{1}{M}\sum_{m=1}^{M}\left[\hat{\eta}_i^m\left(\alpha\right)-\hat{\eta}_i\left(\alpha\right)\right]^2,
$$

and m indexes the simulated samples by parametric bootstrap.

[6] In what follows I omit the argument in ℓ_{it} for notational simplicity.

- Trace Based Approach with Trimming: this approach uses a trimmed version of $\Upsilon_i(\alpha)$, that is,

$$\widehat{\Upsilon}_i(\alpha) = \Omega_0 + 2\sum_{l=1}^{r} \Omega_l,$$

where

$$\Omega_l = \frac{1}{T-l} \sum_{t=l+1}^{T} \left(1 - \frac{l}{r+1}\right) \frac{\partial \ell_{it}}{\partial \eta_i} \cdot \frac{\partial \ell_{it-l}}{\partial \eta_i},$$

for r small. So,

$$\ell_{mi}(\alpha, \hat{\eta}_i(\alpha)) = -\frac{1}{2T} \sum_{t=1}^{T} (y_{it} - \alpha y_{it-1} - \hat{\eta}_i(\alpha))^2 - \frac{1}{2T} \left(\widehat{H}_i^{-1}(\alpha)\, \widehat{\Upsilon}_i(\alpha)\right).$$

- Following Lancaster (2002), I consider the Approximate Conditional Likelihood:

$$\ell_{mi}(\alpha, \hat{\eta}_i(\alpha)) = -\frac{1}{2T} \sum_{t=1}^{T} (y_{it} - \alpha y_{it-1} - \hat{\eta}_i(\alpha))^2 + \frac{b_T(\alpha)}{T},$$

where

$$b_T(\alpha) = \frac{1}{T} \left[\sum_{t=1}^{T-1} \left(\frac{T-t}{t}\right) \alpha^t\right].$$

Before presenting the results I want to mention that I use *Individual Block-Bootstrap,* that is, *fixed-T large-N non parametric bootstrap* for calculating the standard errors of the estimates. The assumption of independence across individual allows me to draw complete time series for each individual to capture the time series dependence, that is, I draw $y_i = (y_{i1}, ..., y_{iT})'$ S times to obtain the simulated data $\left\{y_i^{(s)}, y_{i(-1)}^{(s)}\right\}_{s=1}^{S}$. For each sample I obtain the corresponding estimates of α_0, $\left(\hat{\alpha}^{(1)}, ..., \hat{\alpha}^{(S)}\right)$, and the empirical distribution as an approximation of the distribution of $\hat{\alpha}$.[7]

Table B.1 reports estimates, based on 300 Monte Carlo runs, for $T = 8$ and $N = 500$. I find some differences in the performance between these four types of bias corrections.

---

[7]Notice that, opposite to the block bootstrap procedure used in time-series literature (Hall and Horowitz, 1996; Horowitz, 2003), here I do not need to choose any bandwidth.

I have also found that iterating bias correction, in the case of the first two corrections, improves a bit the estimation but for brevity I do not report here these results. An example of that is included in the next subsection. We see in the table that the fixed effects MLE is downward biased by around 35-40 percent in both cases. Bias corrections, except the one proposed by Lancaster (2002) that is consistent for fixed T, all perform better when $\alpha = 0.5$. In this latter case, the corrections reduce the bias for at least a half. In addition, we can see that the mean of the standard errors estimated by individual block-bootstrap is a good approximation to the Monte Carlo standard deviation.

Table B.2 presents estimates for $T = 16$ and $N = 500$ [8]. We can see that for $\alpha = 0.5$, the MLE has still an important bias, but the modified MLEs are closer to the true value. As before, corrections perform worse when $\alpha = 0.8$.

2.4.2 The linear dynamic panel model with multiple fixed effects

One of the advantages of the bias-correcting estimators with respect to the estimator proposed by Lancaster is their generality. With only a slight modification of the previous expressions it is possible to deal with a more complex model, as an AR(1) model with fixed effects in the conditional mean, η_i, and in the conditional variance, σ_i^2.

Now the model design is

$$\begin{aligned}
y_{it} &= \alpha y_{it-1} + \eta_i + e_{it} = \alpha y_{it-1} + \eta_i + \sigma_i \epsilon_{it}, \;\; (t = 1, ..., T; \; i = 1, ..., N) \\
\epsilon_{it} &\sim N(0,1), \;\; \eta_i \sim N(0,1), \;\; \psi_i = \log \sigma_i^2 \sim N(-3.0, 0.8), \\
y_{i0} &\sim N\left(\frac{\eta_i}{(1-\alpha)}, \frac{\sigma_i^2}{(1-\alpha^2)}\right).
\end{aligned}$$

The data are generated for $T = 8$ and 16, $N = 500$, and for $\alpha = 0.5$. I denote as

[8] I do not report here the results for $N = 1000$, because increasing the number of individuals from $N = 500$ to $N = 1000$ has little effect on the magnitude of the estimated bias (much less effect that increasing T).

$\Theta_i = (\eta_i, \sigma_i^2)'$ the vector of fixed effects. The MLE of α is

$$\begin{aligned}\hat{\alpha} &\equiv \arg\max_{\alpha} \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{T}\sum_{t=1}^{T}\ell_{it}\left(\alpha, \widehat{\Theta}_i(\alpha)\right)\right] \\ &= \arg\max_{\alpha} \frac{1}{N}\sum_{i=1}^{N}\left[-\frac{1}{2}\ln\hat{\sigma}_i^2(\alpha) - \frac{1}{2T}\sum_{t=1}^{T}\frac{(y_{it} - \alpha y_{it-1} - \hat{\eta}_i(\alpha))^2}{\hat{\sigma}_i^2(\alpha)}\right],\end{aligned}$$

where

$$\widehat{\Theta}_i(\alpha) = \begin{pmatrix}\hat{\eta}_i(\alpha) \\ \hat{\sigma}_i^2(\alpha)\end{pmatrix} = \begin{pmatrix}\bar{y}_i - \alpha\bar{y}_{i(-1)} \\ \frac{1}{T}\sum\limits_{t=1}^{T}(y_{it} - \alpha y_{it-1} - (\bar{y}_i - \alpha\bar{x}_i))^2\end{pmatrix},$$

and $\bar{y}_i = \frac{1}{T}\sum\limits_{t=1}^{T} y_{it}$, $\bar{y}_{i(-1)} = \frac{1}{T}\sum\limits_{t=1}^{T} y_{it-1}$, $\tilde{y}_{it} = y_{it} - \bar{y}_i$, $\tilde{y}_{it-1} = y_{it-1} - \bar{y}_{i(-1)}$. Again, I consider several bias-correcting estimators of α that are obtained by maximization of a modified concentrated log likelihood like

$$\widetilde{\alpha} \equiv \arg\max_{\alpha} \frac{1}{N}\sum_{i=1}^{N}\ell_{mi}\left(\alpha, \widehat{\Theta}_i(\alpha)\right).$$

- Determinant Based Approach Using Expected Quantities: now

$$\begin{aligned}H_i(\alpha) &= -\frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}\frac{\partial^2\ell_{it}}{\partial\eta^2} & \frac{\partial^2\ell_{it}}{\partial\eta\partial\sigma^2} \\ \frac{\partial^2\ell_{it}}{\partial\sigma^2\partial\eta} & \frac{\partial^2\ell_{it}}{\partial(\sigma^2)^2}\end{pmatrix} \\ &= \frac{1}{T}\sum_{t=1}^{T}\begin{pmatrix}\frac{1}{\sigma_i^2} & \frac{(y_{it}-\alpha y_{it-1}-\eta_i)}{\sigma_i^4} \\ \frac{(y_{it}-\alpha y_{it-1}-\eta_i)}{\sigma_i^4} & \left(\frac{(y_{it}-\alpha y_{it-1}-\eta_i)^2}{\sigma_i^6}\right) - \frac{1}{2\sigma_i^4}\end{pmatrix},\end{aligned}$$

and

$$\begin{aligned}&\bar{\Upsilon}_i(\alpha, \Theta_i; \alpha_0, \Theta_{i0}) \\ &= TE_0\left\{\left[\frac{\partial\ell_i(\alpha,\Theta_i)}{\partial\Theta_i} - E\left(\frac{\partial\ell_i(\alpha,\Theta_i)}{\partial\Theta_i}\right)\right]\left[\frac{\partial\ell_i(\alpha,\Theta_i)}{\partial\Theta_i'} - E\left(\frac{\partial\ell_i(\alpha,\Theta_i)}{\partial\Theta_i'}\right)\right]\middle| y_{i0}\right\}.\end{aligned}$$

Now, I obtain $\bar{\Upsilon}_i\left(\alpha, \widehat{\Theta}_i(\alpha); \hat{\alpha}, \widehat{\Theta}_i\right)$ as a mean of $\{\Upsilon_i^m(\alpha)\}_{m=1}^M$ calculated in data simulated as $\left\{\prod_{t=1}^T f\left(y_{it}|y_{i0}, \hat{\alpha}, \widehat{\Theta}_i\right)\right\}_{i=1}^N$. That is,

$$\bar{\Upsilon}_i\left(\alpha, \widehat{\Theta}_i(\alpha); \hat{\alpha}, \widehat{\Theta}_i\right) = \frac{1}{M}\sum_{m=1}^M \Upsilon_i^m(\alpha),$$

where

$$\Upsilon_i^m(\alpha) = \frac{1}{T}\sum_{t=1}^T\sum_{s=1}^T\left\{\left[\frac{\partial \ell_{it}}{\partial \Theta_i} - \left(\frac{1}{T}\sum_{r=1}^T \frac{\partial \ell_{ir}}{\partial \Theta_i}\right)\right]\cdot\left[\frac{\partial \ell_{is}}{\partial \Theta_i'} - \left(\frac{1}{T}\sum_{r=1}^T \frac{\partial \ell_{ir}}{\partial \Theta_i'}\right)\right]\right\},$$

and

$$\frac{\partial \ell_{it}}{\partial \Theta_i} = \begin{pmatrix} \frac{\partial \ell_{it}}{\partial \eta} \\ \frac{\partial \ell_{it}}{\partial \sigma^2} \end{pmatrix} = \begin{pmatrix} \frac{(y_{it}-\alpha y_{it-1}-\eta_i)}{\sigma_i^2} \\ \frac{(y_{it}-\alpha y_{it-1}-\eta_i)^2-\sigma_i^2}{2\sigma_i^4} \end{pmatrix}.$$

This leads to

$$\begin{aligned}\ell_{mi}\left(\alpha, \widehat{\Theta}_i(\alpha); \hat{\alpha}, \widehat{\Theta}\right) &= \frac{1}{T}\sum_{t=1}^T \ell_{it}\left(\alpha, \widehat{\Theta}_i(\alpha)\right) + \frac{1}{2T}\ln\det \widehat{H}_i(\alpha) \\ &\quad -\frac{1}{2T}\ln\det \bar{\Upsilon}_i\left(\alpha, \widehat{\Theta}_i(\alpha); \hat{\alpha}, \widehat{\Theta}_i\right).\end{aligned}$$

- Determinant Based Approach Using a Bootstrap Estimate of $Var\left[\widehat{\Theta}_i(\alpha)\right]$: this approach is based on using the bootstrap estimate

$$\widehat{Var}\left[\widehat{\Theta}_i(\alpha)\right] = \frac{1}{M}\sum_{m=1}^M\left[\widehat{\Theta}_i^m(\alpha) - \widehat{\Theta}_i(\alpha)\right]\left[\widehat{\Theta}_i^m(\alpha) - \widehat{\Theta}_i(\alpha)\right]',$$

which leads to

$$\ell_{mi}\left(\alpha, \widehat{\Theta}_i(\alpha)\right) = \frac{1}{T}\sum_{t=1}^T \ell_{it}\left(\alpha, \widehat{\Theta}_i(\alpha)\right) - \frac{1}{2}\ln\det\left(\widehat{H}_i(\alpha)\widehat{Var}\left[\widehat{\Theta}_i(\alpha)\right]\right).$$

- Trace Based Approach with Trimming: this approach uses a trimmed version of $\Upsilon_i(\alpha)$, that is,

$$\widehat{\Upsilon}_i(\alpha) = \Omega_0 + \sum_{l=1}^{r} (\Omega_l + \Omega_l'),$$

where

$$\Omega_l = \frac{1}{T-l} \sum_{t=l+1}^{T} \left(1 - \frac{l}{r+1}\right) \frac{\partial \ell_{it}}{\partial \Theta_i} \cdot \frac{\partial \ell_{it-l}}{\partial \Theta_i'},$$

for r small. So,

$$\ell_{mi}\left(\alpha, \widehat{\Theta}_i(\alpha)\right) = \frac{1}{T} \sum_{t=1}^{T} \ell_{it}\left(\alpha, \widehat{\Theta}_i(\alpha)\right) - \frac{1}{2T}\left(\widehat{H}_i^{-1}(\alpha)\, \widehat{\Upsilon}_i(\alpha)\right).$$

Table B.3 reports estimates for $T = 8$ and 16, and $N = 500$. We see in the table that the fixed effects MLE is downward biased in both cases. Here we can see that iterating bias correction improves substantially the estimation. In fact, bias corrections reduce the bias for at least a half and this bias practically disappears when I iterate the corrections.

2.4.3 The AR(1)-EARCH(1) panel model with fixed effects

Now the model design is

$$\begin{aligned}
y_{it} &= \alpha y_{it-1} + e_{it} = \alpha y_{it-1} + h_{it}^{1/2} \epsilon_{it}, \quad (t = 1, ..., T;\ i = 1, ..., N) \\
h_{it} &= \exp\left(\psi_i + \beta\left[\sqrt{\epsilon_{it-1}^2 + \Lambda} - \sqrt{2/\pi}\right]\right) = h(\epsilon_{it-1}, \psi_i), \\
\epsilon_{it} &\sim N(0,1),\ \psi_i \sim N(-3.0, 0.8).
\end{aligned}$$

where Λ is a small positive number used to approximate the absolute value function by means of a rotated hyperbola, and $\sqrt{2/\pi}$ is an approximation for $E\left[\sqrt{\epsilon_{it-1}^2 + \Lambda}\right]$ given that $\epsilon_{it-1} \sim N(0,1)$. The process is started at $y_{i0} = 0$, then 700 time periods are generated before the sample is generated. I denote as $\Gamma = (\alpha, \beta)$. The data are generated for $T = 8$ and 16, $N = 1000$, $\alpha = 0.5$, and $\beta = 0.5$. For each sample I have estimated Γ by maximum likelihood and, at the moment, by the trimming modified maximum likelihood.

The MLE of Γ is

$$\widehat{\Gamma} \equiv \arg\max_{\Gamma} \frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{T} \sum_{t=1}^{T} \ell_{it} \left(\Gamma, \widehat{\psi}_i \left(\Gamma \right) \right) \right],$$

where

$$\widehat{\psi}_i \left(\Gamma \right) \equiv \arg\max_{\psi} \frac{1}{T} \sum_{t=1}^{T} \ell_{it} \left(\Gamma, \psi \right).$$

Since here I can not get a explicit expression of the fixed effects estimators as functions of α and β, I do a double maximization, strictly speaking N maximizations inside the one for Γ. I use a Quasi-Newton's Method algorithm to maximize the log likelihood function with respect to Γ, and in each step $\widehat{\psi}_i \left(\Gamma \right)$ is computed such that, for this given value of Γ, the individual log likelihood is maximized with respect to ψ.

The MMLE is

$$\begin{aligned} \widetilde{\Gamma} &= \arg\max_{\Gamma} \frac{1}{N} \sum_{i=1}^{N} \ell_{mi} \left(\Gamma, \widehat{\psi}_i \left(\Gamma \right) \right) \\ &= \arg\max_{\Gamma} \frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{T} \sum_{t=1}^{T} \ell_{it} \left(\Gamma, \widehat{\psi}_i \left(\Gamma \right) \right) - \frac{\hat{b}_i \left(\Gamma \right)}{T} \right], \end{aligned}$$

where

$$\hat{b}_i \left(\Gamma \right) = \frac{1}{2} \left[\widehat{H}_i^{-1} \left(\Gamma \right) \widehat{\Upsilon}_i \left(\Gamma \right) \right],$$

for

$$\widehat{H}_i \left(\Gamma \right) = -\frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \ell_{it}}{\partial \psi^2},$$

and a trimmed version of $\Upsilon_i \left(\Gamma \right)$ with r small

$$\widehat{\Upsilon}_i \left(\Gamma \right) = \Omega_0 + 2 \sum_{l=1}^{r} \Omega_l,$$

$$\Omega_l = \frac{1}{T-l}\sum_{t=l+1}^{T}\left(1-\frac{l}{r+1}\right)\frac{\partial \ell_{it}}{\partial \psi_i}\cdot\frac{\partial \ell_{it-l}}{\partial \psi_i}.$$

In this case I calculate numerical first and second derivatives.

Table B.4 reports estimates for $T = 8$ and 16, and $N = 1000$. In this case $\hat{\alpha}$ is not biased, and with the trimming correction I correct an otherwise seriously biased MLE of β.

2.4.4 The AR(1)-EARCH(1) panel model with multiple fixed effects

Here the model design is

$$\begin{aligned}
y_{it} &= \alpha y_{it-1} + \eta_i + e_{it} = \alpha y_{it-1} + \eta_i + h_{it}^{1/2}\epsilon_{it}, \quad (t = 1, ..., T;\ i = 1, ..., N) \\
h_{it} &= \exp\left(\psi_i + \beta\left[\sqrt{\epsilon_{it-1}^2 + \Lambda} - \sqrt{2/\pi}\right]\right) = h\left(\epsilon_{it-1}, \psi_i\right), \\
\epsilon_{it} &\sim N(0,1);\ \eta_i \sim N\left(0,1\right);\ \psi_i \sim N\left(-3.0, 0.8\right).
\end{aligned}$$

The process is started at $y_{i0} = 0$, then 700 time periods are generated before the sample is generated. I denote as $\Gamma = (\alpha, \beta)$. The data are generated for $T = 16$, $N = 1000$, $\alpha_0 = 0.5$, and $\beta_0 = 0.5$. For each sample I have estimated Γ by maximum likelihood and, at the moment, by the trimming modified maximum likelihood.

The MLE of Γ is

$$\widehat{\Gamma} \equiv \arg\max_{\Gamma} \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{T}\sum_{t=1}^{T}\ell_{it}\left(\Gamma, \widehat{\Theta}_i\left(\Gamma\right)\right)\right],$$

where

$$\widehat{\Theta}_i\left(\Gamma\right) \equiv \arg\max_{\Theta} \frac{1}{T}\sum_{t=1}^{T}\ell_{it}\left(\Gamma, \Theta\right),$$

and the MMLE is

$$\begin{aligned}
\widetilde{\Gamma} &= \arg\max_{\Gamma} \frac{1}{N} \sum_{i=1}^{N} \ell_{mi} \left(\Gamma, \widehat{\Theta}_i (\Gamma) \right) \\
&= \arg\max_{\Gamma} \frac{1}{N} \sum_{i=1}^{N} \left[\frac{1}{T} \sum_{t=1}^{T} \ell_{it} \left(\Gamma, \widehat{\Theta}_i (\Gamma) \right) - \frac{\hat{b}_i (\Gamma)}{T} \right],
\end{aligned}$$

where

$$\hat{b}_i (\Gamma) = \frac{1}{2} tr \left[\widehat{H}_i^{-1} (\Gamma) \widehat{\Upsilon}_i (\Gamma) \right],$$

for

$$\widehat{H}_i (\Gamma) = -\frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \ell_{it}}{\partial \Theta \partial \Theta'},$$

and a trimmed version of $\Upsilon_i (\Gamma, \Theta)$

$$\widehat{\Upsilon}_i (\Gamma, \Theta) = \Omega_0 + \sum_{l=1}^{r} (\Omega_l + \Omega_l'),$$

with

$$\Omega_l = \frac{1}{T-l} \sum_{t=l+1}^{T} \left(1 - \frac{l}{r+1} \right) \frac{\partial \ell_{it}}{\partial \Theta} \cdot \frac{\partial \ell_{it-l}}{\partial \Theta'}.$$

Also in this case I calculate numerical first, second and cross derivatives. Table B.5 reports estimates for $T = 16$ and $N = 1000$. Again, I obtain estimates with less bias when I use the modified maximum likelihood estimator.

2.5 Estimation Results

In this section I use the modified maximum likelihood method to estimate an empirical model for the conditional mean and the conditional variance of male wages. As Meghir and Pistafferi (2004), I use data on 2,066 individuals for the period 1968-1993 of the PSID. It is an unbalanced panel with 32,066 observations. I select male heads aged 25 to 55 with at least nine years of usable wages data. Step-by-step details on sample selection are reported in Appendix 2.C and Table B.6. Sample composition by year and by education, and demographic characteristics are presented in Tables B.7-B.9.

The dependent variable is annual real wages of the heads, so I exclude other components of money income for labour as labour part of farm income, business income, overtime, commissions, etc. Figures E.3 and E.4, at the end of the document, plot the mean and the variance of log real wages against time for education group and for the whole sample. These figures look very similar to the ones in Meghir and Pistaferri (2004, pp. 4-5) and, as they say, reproduce well known facts about the distribution of male earnings in the U.S. (Levy and Murnane, 1992).

2.5.1 Estimation of the Model

The dependent variable that I use in the estimation, y_{it}, is log wages residuals from first stage regressions on year dummies, education, a quadratic in age, dummies for race (white), region of residence, and residence in a SMSA[9]. In this version of the model, I deal with aggregate effects in the variance by regarding y_{it} as standardized wages[10].

The equation estimated is

$$y_{it} = \alpha y_{it-1} + \eta_i + e_{it} = \alpha y_{it-1} + \eta_i + \sqrt{h_{it}}\epsilon_{it},\ \ (i = 1, ..., N; t = 0, ..., T)$$

with

$$h_{it} = \exp\left(\psi_i + \beta\left[\sqrt{\epsilon^2_{it-1} + \Lambda} - \sqrt{2/\pi}\right]\right) = h\left(\epsilon_{it-1}, \psi_i\right).$$

Table B.10 presents the estimation results by MLE and by maximization of the

[9]In earnings dynamics research it is standard to adopt a two step procedure. In the first stage regression, the log of real wages is regressed on control variables and year dummies to eliminate group heterogeneities and aggregate time effects. Then, in the second stage, the unobserved heterogeneity and dynamics of the residuals are modelled.

[10]For each year I calculate the sample wage variance and I take $(\log w_{it} - \hat{\mu}_t)/\hat{\sigma}_t$.

trimmed corrected concentrated likelihood. As expected, we can see that the MLE is underestimating the value of α and β. After applying the bias correction, I obtain estimates of both parameters above 0.5. Not only the persistence in the mean is significant. Also the state dependence effects in the volatility of wages seem important.

Correlations between unobserved individual heteregeneity and observed outcomes. One important advantage of the *fixed effects perspective* adopted here is that I also obtain estimates of the unobserved individual heterogeneity and, therefore, I can evaluate the relation between those individual effects in the volatilities of wages and measurable outcomes.

Table B.11 shows that being married, older, and white, are negatively associated with individual fixed effects in the variance. Also, being a technical worker, a manager, or having large values of tenure. On the other hand, being a sales or a services worker, moving from one job to other at least once, or having a low educational degree, are associated with higher volatility. The direction of the association is the one that we could expect.

The $\hat{\psi}_i's$ capture the unobserved heterogeneity in a very robust way. If we were able to observe the individual heterogeneity this would be much better but, if we look at the R^2 of the regression, we can see that with only the observed covariates we can not explain much of the variation across individuals.

Generality of the estimation method. I have also estimated a version of the model similar to Meghir and Windmeijer (1999). It is a convenient specification but more difficult to interpret because the conditional variance of e_{it}, g_{it}, it is a function of the past values of the dependent variable instead of the past values of the error. The model is the following

$$y_{it} = \alpha y_{it-1} + \eta_i + e_{it} = \alpha y_{it-1} + \eta_i + \sqrt{g_{it}}\epsilon_{it};\ (i = 1, ..., N; t = 1, ..., T)$$

with

$$g_{it} = \exp\left(\psi_i + \beta\left[\sqrt{y_{it-1}^2 + \Lambda}\right]\right) = g\left(y_{it-1}, \psi_i\right).$$

Table B.12 presents the corresponding results of the estimation of this model by MLE and by maximization of the trimmed corrected concentrated likelihood. Although the estimates of β are a bit different, the main results do not change.

2.5.2 Checking for Nonnormality

The assumption of normality is not necessary for the validity of the estimation method used on the empirical application, but checking this distributional assumption can be useful for other purposes. The distribution of the errors are nonparametrically identified and can be estimated using deconvolution tecniques as in Horowitz and Markatou (1996). A normal probability plot of residuals in first-differences (Figure E.5) indicates that the tails of the distribution of errors are thicker than those of the normal distribution. However the same plot but for the standardized residuals in first-differences (Figure E.6) is almost a straight line, meaning no deviation from normality[11].

Fit of the model. Given the distributional assumption, parameter estimates, $\hat{\alpha}_T$, $\hat{\beta}_T$, $\hat{\eta}_i$, $\hat{\psi}_i$, and initial conditions, y_{i0}, $\widehat{h_{i1}}$, I simulate an unbalanced panel of standardized logwages observations with the same dimensions as the PSID sample. The first thing that I evaluate with this simulated panel is the fit of the model. Figure E.7 shows the kernel

[11] Estimated residuals and estimated standardized residuals respectively defined as

$$\hat{e}_{it} = y_{it} - \hat{\alpha} y_{it-1} - \hat{\eta}_i.$$

and

$$\hat{\epsilon}_{it} = \frac{y_{it} - \hat{\alpha} y_{it-1} - \hat{\eta}_i}{h_{it}^{1/2}\left(\hat{\psi}_i, \hat{\epsilon}_{it-1}\right)},$$

where

$$h_{it}\left(\hat{\psi}_i, \hat{\epsilon}_{it-1}\right) = \exp\left\{\hat{\psi}_i + \hat{\beta}\left[\left|\hat{\epsilon}_{it-1}\right| - \sqrt{2/\pi}\right]\right\}.$$

densities of logwages in the data and according to the model[12]. It seems that the model performs well.

Individual Heterogeneity. Then, for evaluating the existence of individual heterogeneity on the data, I calculate several counterfactuals in an analogous way. Counterfactual 1 is obtained using the model, the parameter estimates, $\hat{\alpha}_T, \hat{\beta}_T, \hat{\psi}_i$, and initial conditions, $y_{i0}, \widehat{h_{i1}}$, but now $\eta_i = \bar{\eta}, \forall i$, where $\bar{\eta} = N^{-1} \sum_{i=1}^{N} \hat{\eta}_i$. Similarly, counterfactual 2 is obtained using the model, the parameter estimates, $\hat{\alpha}_T, \hat{\beta}_T, \hat{\eta}_i$, and initial conditions, $y_{i0}, \widehat{h_{i1}}$, but now $\psi_i = \bar{\psi}, \forall i$, where $\bar{\psi} = N^{-1} \sum_{i=1}^{N} \hat{\psi}_i$. When I plot the individual means and individual logvariances of logwages (Figures E.8 and E.9, and Table B.13 for some descriptive statistics of those distributions) we can see that there is variation across individuals not only in the means but also in the variances. In addition we can see that the model captures this variation successfully.

Using these counterfactuals I can say how much of the variance in logwages is due to individual heterogeneity in the mean and how much due to individual heterogeneity in the variance according to the model. In particular, for the counterfactual 2, the sample variance of logwages is equal to 0.8581. That is, variation in $\hat{\psi}_i$ accounts for by 14 per cent of the total variation in log wages.

Dynamics: Quantiles of log normal wages. Regarding the dynamics, with a model like the one considered in this chapter I can say how is the effect of lagged values at different parts of the wage distribution. In a general setting, let logwages $y = \log(w) \sim N(\mu, \sigma^2)$ with *cdf*

$$\Pr(\log w \leq r) = \Phi\left(\frac{r - \mu}{\sigma}\right).$$

[12] The bandwidth is equal to 0.10 for all kernels in this section.

The τth quantile of $\log w$, $Q_\tau(\log w)$, is the value of r such that

$$\Phi\left(\frac{Q_\tau(\log w)-\mu}{\sigma}\right)=\tau,$$

so that

$$\frac{Q_\tau(\log w)-\mu}{\sigma}=\Phi^{-1}(\tau)\equiv q_\tau,$$

where q_τ is the τth quantileof the $N(0,1)$ distribution, and

$$Q_\tau(\log w)=\mu+q_\tau\sigma.$$

To get quantiles for w, as opposed to $\log w$, note that

$$\Pr(\log w\leq r)=\Pr(w\leq\exp r),$$

so that

$$\Pr(\log w\leq Q_\tau(\log w))=\Pr(w\leq\exp Q_\tau(\log w))=\tau.$$

Therefore,

$$Q_\tau(w)=\exp Q_\tau(\log w)=\exp(\mu+q_\tau\sigma).$$

Function of $\log w_{it-1}$. In the conditional case, regarding μ and σ as functions of $\log w_{it-1}$,

$$\frac{\partial\log Q_\tau(w_{it})}{\partial\log w_{it-1}}=\frac{\partial\mu}{\partial\log w_{it-1}}+q_\tau\frac{\partial\sigma}{\partial\log w_{it-1}}.$$

In particular, for the model considered here

$$\mu_{it}=\alpha y_{it-1}+\eta_i,$$

$$
\begin{aligned}
\sigma_{it} &= h_{it}\left(\psi_i, \epsilon_{it-1}\right)^{1/2} = \exp\left(\frac{\psi_i}{2} + \frac{\beta}{2}\left[\sqrt{\epsilon_{it-1}^2 + \Lambda} - \sqrt{2/\pi}\right]\right) \\
&= \exp\left(\frac{\psi_i}{2} + \frac{\beta}{2}\left[\sqrt{\left(\frac{y_{it-1} - \alpha y_{it-2} - \eta_i}{h_{it-1}\left(\psi_i, \epsilon_{it-2}\right)^{1/2}}\right)^2 + \Lambda} - \sqrt{2/\pi}\right]\right),
\end{aligned}
$$

and

$$
\frac{\partial \mu_{it}}{\partial y_{it-1}} = \alpha,
$$

$$
\begin{aligned}
\frac{\partial \sigma_{it}}{\partial y_{it-1}} &= \sigma_{it} \times \frac{\beta}{2} \times \frac{1}{\left[\epsilon_{it-1}^2 + \Lambda\right]^{1/2}} \times \left(\frac{y_{it-1} - \alpha y_{it-2} - \eta_i}{h_{it-1}\left(\psi_i, \epsilon_{it-2}\right)^{1/2}}\right) \times \frac{\partial \epsilon_{it-1}}{\partial y_{it-1}} \\
&= \sigma_{it} \times \frac{\beta}{2} \times \frac{\epsilon_{it-1}}{\left[\epsilon_{it-1}^2 + \Lambda\right]^{1/2}} \times \frac{1}{h_{it-1}\left(\psi_i, \epsilon_{it-2}\right)^{1/2}}.
\end{aligned}
$$

Thus I can calculate a mean elasticity at different parts of the wage distribution as

$$
\varepsilon_\tau\left(\log w_{it-1}\right) = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}\left[\frac{\partial \log Q_\tau\left(w_{it}\right)}{\partial \log w_{it-1}}\right].
$$

The first column in Table B.14 shows that those elasticities increase with the quantiles. That is, there are different elasticities below and above the median, where the mean elasticity is just equal to the corrected estimate of alpha, $\hat{\alpha}_T$. In Table B.14 and in Figure E.10, we can see that this pattern is very different for individuals with low (second column) or high (third column) values of the estimated fixed effects in the variance.

Impulse-response function: functions of ϵ_{it-s} Now,

$$
Q_\tau\left(\log w\right) = \mu + q_\tau \sigma.
$$

and in the conditional case, regarding μ and σ as functions of ϵ_{it-1},

$$\frac{\partial Q_\tau (\log w_{it})}{\partial \epsilon_{it-1}} = \frac{\partial \mu}{\partial \epsilon_{it-1}} + q_\tau \frac{\partial \sigma}{\partial \epsilon_{it-1}}.$$

In particular, for the model considered here

$$\begin{aligned}
\mu_{it} &= \alpha y_{it-1} + \eta_i = \alpha \left(\alpha y_{it-2} + \eta_i + h_{it-1} (\psi_i, \epsilon_{it-2})^{1/2} \epsilon_{it-1} \right) + \eta_i \\
\sigma_{it} &= \exp \left(\frac{\psi_i}{2} + \frac{\beta}{2} \left[\sqrt{\epsilon_{it-1}^2 + \Lambda} - \sqrt{2/\pi} \right] \right),
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \mu_{it}}{\partial \epsilon_{it-1}} &= \alpha h_{it-1} (\psi_i, \epsilon_{it-2})^{1/2}, \\
\frac{\partial \sigma_{it}}{\partial y_{it-1}} &= \sigma_{it} \times \frac{\beta}{2} \left(\frac{\epsilon_{it-1}}{\left[\epsilon_{it-1}^2 + \Lambda \right]^{1/2}} \right).
\end{aligned}$$

Thus I can calculate a mean marginal effect at different parts of the logwage distribution as

$$\widehat{E} \left(\frac{\partial Q_\tau (\log w_{it})}{\partial \epsilon_{it-1}} \right) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left[\frac{\partial Q_\tau (\log w_{it})}{\partial \epsilon_{it-1}} \right].$$

Notice that

$$\frac{\partial Q_\tau (\log w_{it})}{\partial \epsilon_{it-1}} = \frac{\partial \log Q_\tau (w_{it})}{\partial \log w_{it-1}} \times \left[h_{it-1} (\psi_i, \epsilon_{it-2})^{1/2} \right].$$

Now, for

$$\frac{\partial Q_\tau (\log w_{it})}{\partial \epsilon_{it-2}} = \frac{\partial \mu}{\partial \epsilon_{it-2}} + q_\tau \frac{\partial \sigma}{\partial \epsilon_{it-2}}.$$

In particular, for the model considered here and

$$
\begin{aligned}
\mu_{it} &= \alpha^2 \left(\alpha y_{it-3} + \eta_i + h_{it-2}\left(\psi_i, \epsilon_{it-3}\right)^{1/2} \epsilon_{it-2}\right) + (1+\alpha)\,\eta_i + \alpha h_{it-1}\left(\psi_i, \epsilon_{it-2}\right)^{1/2} \epsilon_{it-1} \\
&= \alpha^3 y_{it-3} + \left(1+\alpha+\alpha^2\right)\eta_i + \alpha h_{it-1}\left(\psi_i, \epsilon_{it-2}\right)^{1/2} \epsilon_{it-1} + \alpha^2 h_{it-2}\left(\psi_i, \epsilon_{it-3}\right)^{1/2} \epsilon_{it-2}, \\
\sigma_{it} &= \exp\left(\frac{\psi_i}{2} + \frac{\beta}{2}\left[\sqrt{\epsilon_{it-1}^2 + \Lambda} - \sqrt{2/\pi}\right]\right) \\
&= \exp\left(\frac{\psi_i}{2} + \frac{\beta}{2}\left[\sqrt{\left(\frac{y_{it-1} - \alpha y_{it-2} - \eta_i}{h_{it-1}\left(\psi_i, \epsilon_{it-2}\right)^{1/2}}\right)^2 + \Lambda} - \sqrt{2/\pi}\right]\right),
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{\partial \mu_{it}}{\partial \epsilon_{it-2}} &= \alpha^2 h_{it-2}\left(\psi_i, \epsilon_{it-3}\right)^{1/2} + \frac{1}{2}\alpha\beta h_{it-1}\left(\psi_i, \epsilon_{it-2}\right)^{1/2} \frac{\epsilon_{it-1}\epsilon_{it-2}}{\sqrt{\epsilon_{it-2}^2 + \Lambda}}, \\
\frac{\partial \sigma_{it}}{\partial \epsilon_{it-2}} &= \sigma_{it}\frac{\beta}{2}\left(\frac{\epsilon_{it-1}}{\left[\epsilon_{it-1}^2 + \Lambda\right]^{1/2}}\right) \times \sigma_{it-1}\frac{\beta}{2}\left(\frac{\epsilon_{it-2}}{\left[\epsilon_{it-2}^2 + \Lambda\right]^{1/2}}\right).
\end{aligned}
$$

The first panel in Table B.15 shows the mean marginal effects with respect to ϵ_{it-1} over different quantiles of the logwage distribution and the second panel, the case with respect to ϵ_{it-2}. In Figure E.11 we can see that past shocks seem to have effect over logwages even two periods apart.

2.5.3 Job changes

It is important taking into account that in a model where individual heterogeneity is treated as fixed effects we abstract for job changes. A specification like this

$$y_{it} = \alpha y_{it-1} + \eta_i + e_{it},$$

works worse if there are many job changes in the sample because η_i is fixed. In order to evaluate this concern, I consider a sample where individuals in different jobs are treated as different individuals. That is, for each individual

$$
\begin{aligned}
y_{it} &= \alpha y_{it-1} + \eta_{i1} + e_{it}; \text{ individual } i \text{ in job 1,} \\
y_{it} &= \alpha y_{it-1} + \eta_{i2} + e_{it}; \text{ individual } i \text{ in job 2.}
\end{aligned}
$$

I use data on 1,346 and 17,485 observations. I do the same sample selection as before. Sample composition by year and by education, and demographic characteristics are presented in Tables B.16-B.18.

Results are reported in Table B.19. We can see that the significant ARCH effects in the variance disappears as soon as we consider a sample without job changes.

2.5.4 Attrition

A final issue is the extent to which attrition from the PSID has biased the results. In this chapter, I assume that attrition is all accounted for by the permanent characteristics in the individual fixed effects. To provide some evidence for this I compare the estimates in my sample to those obtained using only individuals who are 16 or more years in the sample (921 individuals). This kind of selection mimics attrition bias since it eliminates individuals observed for a shorter time period. The estimates based on this sample are included in Table B.20. The main conclusion is that the corrected estimates are not very different to those reported in Table B.10.

2.6 Implications for Consumption Growth

Given the results above I provide now an example that illustrates the effects that individual risk can have in explaining precautionary saving, that is, additional saving that results from the knowledge that the future is uncertain. Here, I follow most of the literature and I consider that additional saving is achieved by consuming less.

Over the last 30 years there has been a well-documented increase in cross-sectional income inequality in the US, and some authors have suggested that households are now exposed to more earnings instability than they were (Gottschalk and Moffitt, 1994). This figure suggests that precautionary saving motives associated with an increase in income risk could have become more important.

In the presence of complete insurance, either formal or informal, it should only be the component of risk that is common to all individuals in an economy that affects consumption. Banks, Blundell, and Brugiavini (2001) find that it is not the common component of risk, but instead the cohort-specific risks which dominate consumption growth. Their results corroborate the notion that if income uncertainty has been growing over the recent past then the failure of insurance between agents makes the precautionary motive for saving an increasingly important self-insurance mechanism. They use series of repeated cross sections of British households data, but they can not consider individual-specific risk due to the lack of panel data. Here, I evaluate the independent role of individual wage risk in consumption growth.

2.6.1 Consumption Model

Let us consider the following intertemporal consumption model[13] (Browning and Lusardi, 1996), where individuals choose consumption so as to maximize an intertemporal utility function subject to the intertemporal budget constraint:

$$\max_{\{C_{t+k}\}_{k=0}^{T-t}} E_t \sum_{k=0}^{T-t} \left[(1+\delta)^{-k} U\left(C_{t+k}, D_{t+k}\right) \right]$$

$$s.t. \quad A_{t+1+k} = (1 + r_{t+k}) \cdot (A_{t+k} + Y_{t+k} - C_{t+k})$$

$$A_{T+1} \geq 0 \qquad\qquad (k = 0, ..., T-t)$$

[13] I omit the individual index for simplicity.

where, for each period s, C_s is consumption, Y_s labour income or earnings, r_s real interest rate, A_s financial wealth (at the beginning of the period), δ subjective intertemporal rate, and D_s demographic characteristics. I assume the date of death is known and there are not explicit liquidity constraints.

The optimal intertemporal allocation of consumption verifies the Euler equation, that is,

$$E_t\left[\frac{1+r_t}{1+\delta}\cdot\frac{U_C\left(C_{t+1},D_{t+1}\right)}{U_C\left(C_t,D_t\right)}\right]=1$$

where $U_c\left(\cdot\right)$ denotes the first derivative of the utility function with respect to consumption.

I assume a CRRA utility function:

$$U\left(C_t,D_t\right)=\frac{1}{1-\rho}\exp\left(\varphi' D_t\right)\cdot C_t^{1-\rho}$$

where $\rho>0$ is the relative risk aversion coefficient. So,

$$\frac{1+r_t}{1+\delta}\cdot\exp\left(\varphi'\Delta D_{t+1}\right)\cdot\left(\frac{C_{t+1}}{C_t}\right)^{-\rho}=1+\xi_{t+1},$$

where $E_t\left[\xi_{t+1}\right]=0$. Taking logs and using the usual approximation for logs I obtain the linearized Euler equation:

$$\Delta\ln C_{t+1}=\frac{1}{\rho}\ln\left(r_t-\delta\right)+\frac{1}{\rho}\varphi'\Delta D_{t+1}+\frac{1}{2\rho}Var_t\left[\xi_{t+1}\right]+v_{t+1}.$$

The first term on the RHS of the equation takes into account the intertemporal substitution effect: an increase in r_t, opportunity cost of current consumption, implies a higher growth of future consumption. The second term considers how different stages of the life cycle are reflected on the consumption profile, by changes in circumstances implicit in demographic variables. Finally, the third term on the RHS of the equation captures precautionary saving. A rise in the expected variance of earnings innovations represents an increase in earnings risk and should depress period t consumption hence increasing the growth of consumption between t and $t+1$. In other words, a positive parameter implies that risk induces a delay in spending and current consumption is therefore reduced.

Notice that $Var_t\left[\xi_{t+1}\right]$ reflects uncertainty regarding future realizations of any uninsurable variable relevant for consumption. Thus, it is not sufficient to enter the wage risk term alone. A scaling term is required by which "poorer" individuals are more responsive to changes in earnings risk, $\pi_t = \left(\frac{Y_{t-1}}{C_{t-1}}\right)^2$. In consequence,

$$\Delta \ln C_{t+1} = \frac{1}{\rho} \ln\left(r_t - \delta\right) + \frac{1}{\rho}\varphi' \Delta D_{t+1} + \gamma \pi_t \sigma_{t+1}^2 + v_{t+1}$$

where σ_{t+1}^2 is a measure of the conditional variance of the wage shock.

2.6.2 Estimation and results

I use food consumption data from the PSID (1974-1987). In my sample[14], I estimate by OLS[15] the following empirical equation:

$$\Delta \ln C_{it+1} = \delta_t + \beta' \Delta D_{it+1} + \gamma \pi_{it} \sigma_{it+1}^2 + v_{it+1},$$

where σ_{it+1}^2 is replaced by

$$\hat{\sigma}_{it+1}^2 = h_{it+1}\left(\hat{\epsilon}_{it}; \hat{\Gamma}, \hat{\Theta}_i, \text{initial conditions}\right).$$

Looking at the estimate for the γ parameter in Table B.21, column 2, I obtain a

[14]The sample includes 1,191 individuals and 15,192 observations.

[15]It would be interesting to follow the same approach as before considering a complete likelihood function:

$$L_{now} = L_{before} + \sum_{i,t}\left\{-\frac{1}{2}\ln \sigma_v^2 - \frac{1}{2\sigma_v^2}\left[\Delta \ln \hat{C}_{t+1} - \gamma \pi_t \sigma_{t+1}^2\right]^2\right\}$$

where $\Delta \ln \hat{C}_{t+1}$ is obtained from first stage regressions of $\Delta \ln C_{t+1}$ on δ_t and ΔD_{t+1}.

significant and positive effect of this term on the consumption growth. As stated above, an increase in individual risk induces a reduction in current consumption and, therefore, an increase in the growth of consumption between t and $t+1$.

Regarding the interactions with education (columns 3 and 4), we can see that this positive effect is more important for the less educated people, slightly significant for the graduate and insignificant for the college educated. This result goes in line with the idea that there are more insurance possibilities for these latter.

2.7 Conclusions

In this chapter I propose a model for the conditional mean and the conditional variance of individual wages. It is a non linear dynamic panel data model with multiple individual fixed effects. For estimating the parameters of the model I assume a distribution for the shocks and apply bias corrections to the concentrated likelihood. This corrects the bias of the estimated parameters from $O\left(T^{-1}\right)$ to $O\left(T^{-2}\right)$, so the estimator has a good finite sample performance and a reasonable asymptotic approximation for moderate T. In fact, Monte Carlo results show that the bias of the MLE is substantially corrected for samples designs that are broadly calibrated to the PSID dataset.

The main advantage of this approach is its generality. As we have seen, the method is generally applicable to take into account dynamics and multiple fixed effects. Another advantage is that the fixed effects are estimated as part of the estimation process.

The empirical analysis is conducted on data drawn from the 1968-1993 PSID dataset. In line with previous literature, I find a corrected estimate for the autoregressive coefficient in the mean around 0.5 (Alvarez and Arellano, 2004), and positive ARCH effects for the variance (Meghir and Pistafferri, 2004). Job changes are driving this dynamics in the variance. I also find important fixed differences across individuals in the variance. In addition, it turns out that this located-scaled model explains the non-normality observed in logwage data. I then illustrate some implications that ARCH effects may have in the field of savings.

Finally there are three issues, at least, that require further research: measurement error in PSID wages, a more comprehensive model that include job changes, and the comparison with female workers in terms of wage profiles.

Appendix of Chapter 2

2.A Bias of the Concentrated Likelihood

Following Arellano and Hahn (2006a, 2006b), let us obtain the expression for the First Order Bias of the Concentrated Likelihood at an arbitrary value of the common parameter Γ. Let $\ell_i(\Gamma, \Theta_i) = \sum_{t=1}^{T} \ell_{it}(\Gamma, \Theta_i)/T$ where $\ell_{it}(\Gamma, \Theta_i) = \ln f(y_{it}|y_{it-1}, \Gamma, \Theta_i)$ denotes the log likelihood of one observation. Let

$$\overline{\Theta}_i(\Gamma) = \arg\max_{\Theta_i} \operatorname{plim}_{T\to\infty} \ell_i(\Gamma, \Theta_i),$$

and

$$\widehat{\Theta}_i(\Gamma) = \arg\max_{\Theta_i} \ell_i(\Gamma, \Theta_i),$$

so that under regularity conditions $\overline{\Theta}_i(\Gamma_0) = \Theta_{i0}$.

Following Severini (2000) and Pace and Salvan (2005), the concentrated likelihood for unit i

$$\hat{\ell}_i(\Gamma) = \ell_i\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right),$$

can be regarded as an estimate of the unfeasible concentrated log likelihood

$$\bar{\ell}_i(\Gamma) = \ell_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right).$$

Now, define

$$u_{it}(\Gamma,\Theta_i)=\frac{\partial \ell_{it}(\Gamma,\Theta_i)}{\partial \Gamma},\ v_{it}(\Gamma,\Theta_i)=\frac{\partial \ell_{it}(\Gamma,\Theta_i)}{\partial \Theta_i},$$

$$u_i(\Gamma,\Theta_i)=\frac{1}{T}\sum_{t=1}^{T}u_{it}(\Gamma,\Theta_i),\ v_i(\Gamma,\Theta_i)=\frac{1}{T}\sum_{t=1}^{T}v_{it}(\Gamma,\Theta_i),$$

$$H_i(\Gamma)=-\lim_{T\to\infty}E\left[\frac{\partial v_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)}{\partial \Theta_i'}\right].$$

When Θ_{i0} is a vector of fixed effects, the Nagar expansion for $\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)$ takes the form

$$\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)=H_i^{-1}(\Gamma)\,v_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)+\frac{1}{T}B_i(\Gamma)+O_p\left(\frac{1}{T^{3/2}}\right), \tag{A.1}$$

where

$$\begin{aligned} B_i(\Gamma) \;=\;& H_i^{-1}(\Gamma)\Big[\Xi_i(\Gamma)\,vec\left(H_i^{-1}(\Gamma)\right) \\ &+\frac{1}{2}E\left(\frac{\partial}{\partial\Theta'}vec\frac{\partial v_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)}{\partial\Theta'}\right)'\left(H_i^{-1}(\Gamma)\otimes H_i^{-1}(\Gamma)\right)vec\left(\Upsilon_i(\Gamma)\right)\Big], \end{aligned}$$

and

$$\Upsilon_i(\Gamma)=\Upsilon_i(\Gamma;\Gamma_0,\Theta_{i0})=\lim_{T\to\infty}TE\left[v_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)v_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)'\right],$$

$$\Xi_i(\Gamma)=\Xi_i(\Gamma;\Gamma_0,\Theta_{i0})=\lim_{T\to\infty}TE\left[\frac{\partial v_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)}{\partial\Theta'}\otimes v_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)'\right].$$

Next, expanding $\ell_i\left(\Gamma,\widehat{\Theta}_i(\Gamma)\right)$ around $\overline{\Theta}_i(\Gamma)$ for fixed Γ,

$$\begin{aligned} & \ell_i\left(\Gamma,\widehat{\Theta}_i(\Gamma)\right)-\ell_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right) \\ =\;& \frac{\partial \ell_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)}{\partial\Theta'}\left(\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right) \\ &+\frac{1}{2}\left(\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right)'\frac{\partial^2\ell_i\left(\Gamma,\overline{\Theta}_i(\Gamma)\right)}{\partial\Theta\partial\Theta'}\left(\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right)+O_p\left(\frac{1}{T^{3/2}}\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{\partial \ell_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)}{\partial \Theta'}\left(\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right) \\
&\quad +\frac{1}{2}\left(\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right)' E\left(\frac{\partial^2 \ell_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)}{\partial \Theta \partial \Theta'}\right)\left(\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right)+O_p\left(\frac{1}{T^{3/2}}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= v_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)'\left(\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right) \\
&\quad -\frac{1}{2}\left(\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right)' H_i(\Gamma)\left(\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right)+O_p\left(\frac{1}{T^{3/2}}\right).
\end{aligned}
$$

Substituting (A.1)

$$
\ell_i\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right)-\ell_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)=\frac{1}{2} v_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)' H_i^{-1}(\Gamma) v_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)+O_p\left(\frac{1}{T^{3/2}}\right).
$$

Taking expectations

$$
E\left[\ell_i\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right)-\ell_i\left(\Gamma, \overline{\Theta}_i(\Gamma)\right)\right]=\frac{1}{2T} tr\left(H_i^{-1}(\Gamma) \Upsilon_i(\Gamma)\right)+O_p\left(\frac{1}{T^{3/2}}\right).
$$

So the bias in the expected concentrated likelihood at an arbitrary Γ is

$$
b_i(\Gamma)=\frac{1}{2} tr\left(H_i^{-1}(\Gamma) \Upsilon_i(\Gamma)\right)=\frac{1}{2} tr\left(H_i(\Gamma) Var\left(\sqrt{T}\left[\widehat{\Theta}_i(\Gamma)-\overline{\Theta}_i(\Gamma)\right]\right)\right).
$$

Thus,

$$
\sum_{i=1}^N \sum_{t=1}^T \ell_{it}\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right)-\sum_{i=1}^N \widehat{b}_i(\Gamma),
$$

is expected to be a closer approximation to the target likelihood than $\sum_{i=1}^N \sum_{t=1}^T \ell_{it}\left(\Gamma, \widehat{\Theta}_i(\Gamma)\right)$.

Moreover, in the likelihood context, it is appropriate to consider a local version of the estimated bias (Pace and Salvan 2005) constructed as an expansion of $\widehat{b}_i(\Gamma)$ at Γ_0 using that at the truth

$$H_i^{-1}\left(\Gamma_0\right)\Upsilon_i\left(\Gamma_0\right) = 1.$$

Taking $\widehat{b}_i\left(\Gamma\right) = \frac{1}{2}tr\left(\widehat{H}_i^{-1}\left(\Gamma\right)\widehat{\Upsilon}_i\left(\Gamma\right)\right)$ also

$$\widehat{b}_i\left(\Gamma\right) = \frac{1}{2}p + \frac{1}{2}\sum_{j=1}^{p}\left[\lambda_j\left(\widehat{H}_i^{-1}\left(\Gamma\right)\widehat{\Upsilon}_i\left(\Gamma\right)\right) - 1\right],$$

where $\lambda_j\left(\widehat{H}_i^{-1}\left(\Gamma\right)\widehat{\Upsilon}_i\left(\Gamma\right)\right)$ denotes the j-th eigenvalue of $\widehat{H}_i^{-1}\left(\Gamma\right)\widehat{\Upsilon}_i\left(\Gamma\right)$ and p is the dimension of Γ. Thus a local version of $\widehat{b}_i\left(\Gamma\right)$ gives

$$\widehat{b}_i\left(\Gamma\right) = \frac{1}{2}p + \frac{1}{2}\sum_{j=1}^{p}\left[\lambda_j\left(\widehat{H}_i^{-1}\left(\Gamma\right)\widehat{\Upsilon}_i\left(\Gamma\right)\right)\right] + O_p\left(\frac{1}{T}\right).$$

Moreover

$$\begin{aligned}\frac{1}{2}\sum_{j=1}^{p}\left[\lambda_j\left(\widehat{H}_i^{-1}\left(\Gamma\right)\widehat{\Upsilon}_i\left(\Gamma\right)\right)\right] &= \frac{1}{2}\ln\det\left(\widehat{H}_i^{-1}\left(\Gamma\right)\widehat{\Upsilon}_i\left(\Gamma\right)\right) \\ &= -\frac{1}{2}\ln\det\widehat{H}_i\left(\Gamma\right) + \frac{1}{2}\ln\det\widehat{\Upsilon}_i\left(\Gamma\right),\end{aligned}$$

which provided justification for the bias-corrected concentrated that I have used.

2.B Analytical expression for $\bar{\Upsilon}_i\left(\alpha, \widehat{\eta}_i\left(\alpha\right); \hat{\alpha}, \hat{\eta}_i\right)$ in the AR(1) model

Let us obtain an expression for $\bar{\Upsilon}_i\left(\alpha, \widehat{\eta}_i\left(\alpha\right); \hat{\alpha}, \hat{\eta}_i\right)$ in the dynamic panel example:

$$y_{it} = \alpha y_{it-1} + \eta_i + \epsilon_{it},$$

where $\epsilon_{it} \sim iidN\left(0,1\right)$. Then

$$\begin{aligned}
\ell_{it}\left(\alpha,\eta\right) &= C-\frac{1}{2}\left(y_{it}-\alpha y_{it-1}-\eta_i\right)^2, \\
\frac{\partial \ell_{it}\left(\alpha,\eta\right)}{\partial \eta} &= y_{it}-\alpha y_{it-1}-\eta_i \equiv v_{it}\left(\alpha,\eta\right) \equiv v_{it}, \\
\bar{v}_i &= \frac{1}{T}\sum_{t=1}^{T} v_{it},
\end{aligned}$$

and

$$\bar{\Upsilon}_i\left(\alpha,\eta;\alpha_0,\eta_0\right) = TVar_0\left(\overline{v}_i|y_{i0}\right).$$

Note that

$$\begin{aligned}
v_{it} &= \epsilon_{it}+\left(\alpha_0-\alpha\right)y_{it-1}+\left(\eta_{i0}-\eta_i\right), \\
\bar{v}_i &= \bar{\epsilon}_i+\left(\alpha_0-\alpha\right)\bar{y}_{i(-1)}+\left(\eta_{i0}-\eta_i\right),
\end{aligned}$$

$$Var_0\left(\bar{v}_i|y_{i0}\right) = \frac{1}{T}+\left(\alpha_0-\alpha\right)^2 Var_0\left(\bar{y}_{i(-1)}|y_{i0}\right)+2\left(\alpha_0-\alpha\right)Cov_0\left(\bar{y}_{i(-1)},\bar{\epsilon}_i|y_{i0}\right),$$

where $\bar{y}_{i(-1)} = \frac{1}{T}\sum\limits_{t=1}^{T} y_{it-1}$. Since

$$\begin{aligned}
\bar{y}_{i(-1)} &= h_T\left(\alpha_0\right)\eta_{i0}+c_T\left(\alpha_0\right)y_{i0}+ \\
&\quad \frac{1}{T}\left[\left(1+\alpha_0+\ldots+\alpha_0^{T-2}\right)\epsilon_{i1}+\left(1+\alpha_0+\ldots+\alpha_0^{T-3}\right)\epsilon_{i2}+\ldots+\epsilon_{iT-1}\right],
\end{aligned}$$

where

$$\begin{aligned}
h_T\left(\alpha_0\right) &= \frac{1}{T}\left[1+\left(1+\alpha_0\right)+\left(1+\alpha_0+\alpha_0^2\right)+\ldots+\left(1+\alpha_0+\ldots+\alpha_0^{T-2}\right)\right], \\
c_T\left(\alpha_0\right) &= \frac{1}{T}\left(1+\alpha_0+\ldots+\alpha_0^{T-1}\right).
\end{aligned}$$

Thus

$$Var_0\left(\bar{y}_{i(-1)}|y_{i0}\right) = \frac{1}{T^2}\left[1 + \left(1+\alpha_0\right)^2 + \ldots + \left(1+\alpha_0+\ldots+\alpha_0^{T-2}\right)^2\right] \equiv \omega_T\left(\alpha_0\right),$$

$$Cov_0\left(\bar{y}_{i(-1)}, \bar{\epsilon}_i|y_{i0}\right) = \frac{1}{T^2}\left[\left(1+\alpha_0+\ldots+\alpha_0^{T-2}\right)+\ldots+1\right] \equiv \psi_T\left(\alpha_0\right),$$

$$Var_0\left(\bar{v}_i|y_{i0}\right) = \frac{1}{T} + \left(\alpha_0-\alpha\right)^2\omega_T\left(\alpha_0\right) + 2\left(\alpha_0-\alpha\right)\psi_T\left(\alpha_0\right),$$

and

$$\bar{\Upsilon}_i\left(\alpha,\eta;\alpha_0,\eta_0\right) = 1 + T\left(\alpha_0-\alpha\right)^2\omega_T\left(\alpha_0\right) + 2T\left(\alpha_0-\alpha\right)\psi_T\left(\alpha_0\right).$$

Thus

$$\bar{\Upsilon}_i\left(\alpha,\widehat{\eta}_i\left(\alpha\right);\hat{\alpha},\hat{\eta}_i\right) = 1 + T\left(\hat{\alpha}-\alpha\right)^2\omega_T\left(\hat{\alpha}\right) + 2T\left(\hat{\alpha}-\alpha\right)\psi_T\left(\hat{\alpha}\right).$$

2.C Sample Selection

Starting point: PSID 1968-1993 Family and Individual - merged files (53,005 individuals).

1. Drop members of the Latino sample (10,022 individuals) and those who are never heads of their households (26,945 individuals).

   = Sample (16,038 individuals)

2. Keep only those who are continuously heads of their households, keep only those who are in the sample for 9 years or more, and keep only those aged 25 to 55 over the period.

   = Sample (5,247 individuals)

3. Drop female heads.

   = Sample (4,036 individuals)

4. Drop those with a spell of self-employment, drop those with missing earnings, and drop those with zero or top-coded earnings data.

   = Sample (2,205 individuals)

5. Drop those with missing education and race records, and those with inconsistent education records.

   = Sample (2,148 individuals)

6. Drop those with outlying earnings records, that is, a change in log earnings greater than 5 or less than -3 and those with noncontinuous data.

   = FINAL SAMPLE (2,066 individuals and 32,066 observations).

Chapter 3

Job Changes and Individual-Job-Specific Wage Dynamics

3.1 Introduction

A large literature on labour economics has focused on the determinants of wages. On the one hand, studies based on the human capital theory (Becker, 1975) examine the impact of general experience on wages, ignoring job mobility. On the other hand, studies based on job search and matching theories (Burdett, 1978; Jovanovic 1979) or purely learning by doing (Rosen, 1972), look at the effect of job specific human capital. This literature has focused on estimating the returns of experience and tenure[1], trying to control for the endogeneity of tenure using different methods[2].

Another related literature on earnings dynamics have modelled and estimated the

[1]See, for example, Altonji and Shakotko (1987), Topel (1991), Topel and Ward (1992), Neal (1995), Altonji and Williams (1997), Dustmann and Meghir (2005), among others.

[2]A group of studies uses a single wage equation and then apply instrument variable or control function methods to control for the endogeneity bias (Altonji and Shakotko, 1987; Topel, 1991; Altonji and Williams, 1997; Dustmann and Meghir, 2005). A second approach is to exploit information on firm closures (Neal 1995, Bonhomme and Jolivet, 2006). A third group suppose that workers' mobility decisions produce realized wage rates that are not random samples of the offered wage rates and estimate the returns of tenure taking into account the sample selection process (Topel, 1986; Marshall and Zarkin, 1987). Finally, other studies explicitly specifies a simultaneous equation model with the wage rate and job tenure as dependent variables, based upon a model in which they are jointly determined (Lillard, 1999; Abowd and Kang, 2002; Bagger, 2007; Amann and Klein, 2007).

heterogeneity and time series properties of individual wage processes (Lillard and Willis, 1978; MaCurdy, 1982; Abowd and Card (1989), among others), but many have ignored job mobility and the distinction between dynamics within and between jobs. In the second chapter of this thesis, I consider a model for the heterogeneity and dynamics of the conditional mean and the conditional variance of individual wages. In the empirical analysis of that chapter - conducted on data drawn from the 1968-1993 Panel Study of Income Dynamics (PSID) - I find that it is important to account for individual unobserved heterogeneity and dynamics also in the conditional variance, and that the latter is driven by job mobility. In line with those results, this chapter develops a model that explicitly considers job changes in the dynamics of wages and in the heterogeneity pattern. In particular, the specification proposed has two different parameters to capture dynamics within jobs and across jobs, and the unobserved heterogeneity shows a richer pattern, as well, composed of both individual and job-specific effects.

As pointed out by Low *et al.* (2007), it is important to distinguish between movements in earnings that reflect choice and those which reflect uncertainty. Those authors address this issue by allowing for endogenous labour supply and job mobility which implies that a proportion of earnings fluctuations, usually interpreted as risk, are in fact attributed to choice. Here, the potential endogeneity of job mobility in relation to earnings is circumvented using an instrument variable estimation method that controls for individual and job-specific unobserved heterogeneity. This match effect will change across jobs but it will remain constant within a position[3]. Differently to Lillard (1999), Abowd and Kang (2002) and Low *et al.* (2007), I adopt a fixed effects perspective leaving the distribution for the unobserved heterogeneity components completely unrestricted and treating each effect as one different parameter to be estimated.

In the empirical application, I use data on work histories drawn from the PSID, which

[3]The importance of match effects in explaining wages has been stressed by Topel and Ward (1992), Abowd, Kramarz and Margolis (1999), Postel-Vinay and Robin (2002) and Bonhomme and Jolivet (2006).

allows the distinction between voluntary and involuntary job-to-job changes. In the data, once we control for individual and job-specific effects, the dynamics within jobs is almost zero, whereas across jobs is significant but small. For the dynamics, the distinction between voluntary and involuntary transitions turns out to be irrelevant. However, that distinction matters in the case of the components of the cross-sectional variance. The estimated variance of the job-specific effects represents around one third of the variance for the individual fixed effects. If I consider a subsample that only includes involuntary job changes, the estimated variance of the heterogeneity across jobs increases up to one half.

This chapter contributes to the literature by more thoroughly describing the impact of job mobility on the dynamics and heterogeneity of individual wages than previous references. First, the model permits that job changes may be correlated with individual and job specific characteristics. Second, it is agnostic regarding the distribution of these individual and job effects. Third, it can be estimated with no need to explicitly model the job mobility process. Finally, the model also allows calculating different components of variance within and between jobs.

The rest of the chapter is organised as follows. Section 3.2 describes the data. Section 3.3 presents the model. Section 3.4 explains the estimation strategy and section 3.5 shows the estimation results. Finally, section 3.6 concludes with a future research agenda.

3.2 The Data

The data come from the PSID for the period 1968-1993. The PSID began in 1968 by interviewing over 5,000 families. Of these, about 3,000 families were representative of the US population as a whole (the core sample), and about 2,000 were low-income families (the Census Bureaus SEO sample). Thereafter, these same families have subsequently been interviewed every year, as have any new families formed from the original group of families[4]. The survey contains abundant information on individual characteristics, income and labour market status. The data set should follow individuals over a sufficiently long period of time to observe pre- and post- job changes earnings histories.

3.2.1 Sample Construction

In the empirical analysis, I use the core sample. I restrict my study to heads of households since survey questions on the PSID regarding employment history are only asked to household heads[5] and, only from 1979, also to wives. In addition, I select males aged 25 to 55 - to focus the analysis during the working life - with no missing records on race, education or region of residence. I drop those with top coded wages, the self-employed, those with less than 8 years of usable data on earnings and those with missing records on the question reason of change. Finally, I have an unbalanced panel that contains 2,013 individuals and 27,845 observations from 1968 to 1992[6]. Step-by-step details on sample selection are reported in Appendix 3.A, and sample composition by year, individuals by number of observations and demographic characteristics are presented in Tables C.1-C.3.

3.2.2 Job Changes Definition

I determine that a job change takes place if current tenure of the worker is less than a year and if there is information available regarding the type of job change. The type of change is defined by the answer to the question, "What happened to the job you had before - did the company go out of business, were you laid off, promoted, or what?". That question

[4] A family member who moves out of a PSID family is eligible for interviewing as a separate family unit if he or she is a sample member and he or she is 18 years old or older and living in a different, independent household.

[5] A household head is defined as the adult of the family. When there is more than one adult in the family, the PSID assigns the primary male adult as the household head.

[6] Since time reference for wage records is the previous year in every survey wave, I use information only until 1992.

was only asked to individuals who report being with their present employer for less than twelve months (otherwise the question is skipped and coded as not applicable), so this make me feel confident regarding the variable tenure[7]. As pointed out by Polsky (1999) from 1984-88, this question was asked of all respondents who reported that their current job started after January of the previous year. To correct for this possible inconsistency, no job change is reported for those with current tenure greater than one year.

From the answers to the question regarding reason of change, I define a job change as an involuntary job separation or job loss in case of business or plant closing or due to being laid off or fired; and as quit, in case of voluntary change.

The sample only includes job-to-job changes, because monthly calendar information (that would provide information regarding spells of unemployment with durations of less than a year) is not available in the PSID prior 1984.

3.2.3 Descriptive Analysis of the Raw Data

The descriptive analysis will emphasize a number of salient facts about job mobility and the relationship between this and earnings dynamics.

Job mobility Among the 2,013 sample individuals, there 699 individuals (around 35 percent) who never change job, whereas the remaining individuals change at least once (on average they have 3.40 different jobs).

As pointed out by Topel and Ward (1992), the most prominent and widely documented facts about job mobility are that average rates of job changing decline with age or experience and, specially, with current job tenure. These facts are consistent with the

[7] Because the PSID did not collect information on specific employers, the identification of job changes in this data set has been quite controversial. Many of the difficulties related to measuring job tenure in the PSID were evaluated by Brown and Light (1992). The tenure question also switched from being coded in intervals prior to 1976 to being measured in months, and from asking about "position" tenure to "employer" tenure. In any case, these difficulties diminish here since I am not interested in the exact value of the variable but if it is less or more than one year.

predictions of job-matching and search models[8] (Johnson, 1978; Burdett, 1978; Jovanovic, 1979). Figure E.12 shows those patterns in the sample.

Regarding "vintage effects", it is less clear if people entering the labour market more recently have patterns of labour mobility different from those of earlier cohorts. Table C.4 presents the distribution of jobs by birth cohort. The 1921-1941 cohort contains a larger proportion of individuals who only have one job than individuals born between 1941-1960. Although sample selection may be relevant, since workers are more exposed to job changes as they grow older and more recent entrants are less likely to be observed in higher-order jobs, the results in the table suggest an increase in job instability for the most recent cohort in the sample.

With respect to the job-exit reason, if we look (Figure E.13) at average rates of job changing by cohorts we find that younger cohorts of workers are more likely to be laid off from their jobs than older cohorts but the difference is bigger in case of quit. More striking is the comparison across skill groups. For all groups the main reason for leaving job is quitting, but the difference with respect to layoff is more important for graduate and - specially - for college people than for dropouts.

Job mobility and earnings dynamics In order to get a first impression of the impact that job changes have over the evolution of earnings (and as a check of the definitions above), I calculate the cross-sectional sample correlations for consecutive logwage observations on years when no-change, a job loss or a job quit has happened. I deflate nominal annual earnings by the GNP Personal Consumption Expenditure Deflator (base 1992). Table C.5 summarizes those calculations. As we would expect, when a job change occurs the correlation diminish, and that reduction is bigger in case of loss than in case

[8]In a matching model, job mobility is the consequence of a voluntary change to a better position where the worker is more productive and receives a higher pay. Search models are based in the existence of imperfect information. In these models, jobs are experience goods. As time goes by, the firm acquires more information and it can adjust the salary better. Under this approach, job mobility is the result of a "poor" matching looking for a better chance.

of a voluntary change.

Table C.6 displays average annual wage growth for workers within jobs and between jobs by type of exit. Within job average annual wage growth is lower than between job average wage growth in case of voluntary transitions. In case of job loss I obtain a drop on real wages. I find the same qualitative pattern among different demographic groups.

As pointed out by Dustmann and Meghir (2005), the fact that within job average annual wage growth is lower than between job average wage growth does not imply that, on average, job quitters have higher wages than stayers. As they did, I regress log wages on dummies for the number of jobs workers have held up to then, also including age and year dummies. Estimates for the first seven jobs, reported in the first column of Table C.7, indicate that workers with more jobs have lower wages. Once I include individual fixed effects in the regression (column 4), the number of jobs is positively related with wages. In fact, if I exclude from the sample movers who transit only through job loss (columns 2 and 5), I obtain a positive relationship between number of jobs and wages. On the contrary, if I exclude those who change voluntary (columns 3 and 6), I obtain that workers with more jobs have lower wages even after including individual fixed effects.

3.3 The Model

In this section I propose an empirical model to study the dynamics of individual earnings over time, within a job and over the career of a worker in one or more different jobs.

3.3.1 Basic Specification

Building on the autoregressive model developed in Lillard and Willis (1978), for a worker i that is observed for T_i periods always at the same job, I consider the following standard specification

$$y_{it} = \alpha y_{it-1} + v_{it} = \alpha y_{it-1} + \eta_i + \epsilon_{it}; \ (t = 1, .., T_i), \ \ |\alpha| < 1,$$

where y_{it} is the log earnings of an individual i in period t, the parameter α measures the persistence on the level of those earnings to shocks, η_i is an unobserved time invariant individual component, like ability, and ϵ_{it} is a purely transitory person-period component, that is, $E\left(\epsilon_{it} | y_i^{t-1}, \eta_i\right) = 0$, where $y_i^{t-1} = (y_{i1}, ..., y_{it-1})'$.[9] I abstract from additive aggregate effects by regarding y_{it} as a deviation from a time effect[10].

Given the model and the initial condition, y_{i1}, the wage profile of an individual i who always stays at the same job would evolve as

$$\begin{aligned} y_{i2} &= \alpha y_{i1} + v_{i2} = \alpha y_{i1} + \eta_i + \epsilon_{i2} \\ y_{i3} &= \alpha y_{i2} + v_{i3} = \alpha y_{i2} + \eta_i + \epsilon_{i3} \\ &\vdots \end{aligned}$$

that is, her wage today would be α times her wage yesterday (where the parameter α measures the persistence on the level of wages to shocks) plus a random term, v_{it}, due to

---

[9] In the sequel, for any random variable (or vector of variables) Z, z_{it} denotes observation for individual i at period t, and $z_i^t = \{z_{i1}, ..., z_{it}\}$, i.e. the set of observations for individual i from the first period to period t.

[10] As is usual in the earnings dynamics literature, the variable y_{it} - strictly speaking - represents log earnings residuals from first stage regressions on some observed variables -apart from year dummies (that capture the aggregate conditions of the economy) - as age, race and other individual characteristics. So we would keep in mind the following structure:

$$\begin{aligned} w_{it} &= x_{it}\beta + u_{it} \\ u_{it} &= \gamma_i + v_{it} \\ v_{it} &= \alpha v_{it-1} + \epsilon_{it} \end{aligned}$$

where w_{it} is the log annual wages of an individual i in period t, x_{it} is a vector of exogenous variables, and u_{it} is a random error with two components , an unobserved individual heterogeneity component and an autoregressive component. The connection with the specification proposed above would be $y_{it} = \hat{u}_{it}$ and $\eta_i = (1 - \alpha)\gamma_i$.

an unobserved time invariant individual component, η_i, like ability, and a purely transitory person-period component ϵ_{it}.

On the contrary, for a worker h that changes job between $t = 3$ and $t = 4$, I would consider

$$y_{h2} = \alpha y_{h1} + \eta_h + \epsilon_{h2}$$

$$\left.\begin{array}{r} y_{h3} = \alpha y_{h2} + \eta_h + \epsilon_{h3} \\ y_{h4} = \alpha^* y_{h3} + \eta_h^* + \epsilon_{h4} \end{array}\right\} \begin{array}{l} \text{job change} \Rightarrow h \text{ ends job at } t{=}3 \\ \text{and starts a new one at } t{=}4 \end{array}$$

$$y_{h5} = \alpha y_{h4} + \eta_h^* + \epsilon_{h5}$$

$$\vdots$$

This specification departs from the standard one in two main features related to job mobility:

1. The dynamics captured by the autoregressive parameters is different in years when workers change job, $\alpha^* = \alpha + \beta$, than within the same job, α.

2. The unobserved individual heterogeneity have a job-specific matching component. In other words, I consider individual and job specific fixed effects, $\eta_{i(t)} = \mu_i + \phi_{ij}$; that is, within the same job we would have $\eta_{i(t)} = \eta_{i(t-1)}$, but between jobs $\eta_{i(t)} = \mu_i + \phi_{ij}^* \neq \eta_{i(t-1)} = \mu_i + \phi_{ij}$. As mentioned before, I adopt a fixed effects perspective leaving the distribution for the unobserved heterogeneity completely unrestricted both within jobs as well as between jobs.

To sum up, the general formulation of the model is the following

$$\begin{aligned} y_{it} &= \alpha y_{it-1} + \beta d_{it-1} y_{it-1} + v_{it} = \alpha y_{it-1} + \beta d_{it-1} y_{it-1} + \eta_{i(t)} + \epsilon_{it} \\ &= \alpha y_{it-1} + \beta d_{it-1} y_{it-1} + \mu_i + \phi_{ij} + \epsilon_{it};\ (i = 1, ...N; t = 2, .., T_i), \end{aligned} \tag{3.1}$$

where d_{it} is an indicator of worker i ending current job at time t.[11]

Given the model, within job, the transitory shocks will be uncorrelated with lagged earnings, but not with present or future earnings. Similarly, I do not need to assume the strict exogeneity of the job changes, in the sense of being uncorrelated to past, present, and future time-varying shocks. Apart from possibly being correlated with the unobserved heterogeneity components, I will consider that job changes may be predetermined, that is, they might be correlated with errors at certain periods but not at others. In particular, we could think on d_{it} as a function of past errors, $d_{it} = f\left(\epsilon_{it-1}, \epsilon_{it-2}, \epsilon_{it-3}, ...\right)$, and unobserved heterogeneity components - that is, the individual's work history - but as being uncorrelated to present and future shocks. Specifically, I am imposing that

$$E\left(\epsilon_{it}|y_i^{t-1}, d_i^t\right) = 0. \tag{3.2}$$

Although it would be preferable to also allow for correlation between d_{it} and ϵ_{it}, that would lead us to consider selection models which is out of the scope of this thesis. Even so, this specification has several advantages. First, it permits the estimation of a model in which job changes can be correlated with individual and job specific characteristics. Second, I do not need to do any assumption regarding the distribution of these individual and job effects. Third, I do not need either to explicitly model the job mobility process. The model also allows to calculate different components of variance within and between jobs. Moreover, note that neither time series nor conditional heteroskedasticity are assumed. That is, the unconditional variances of the errors, denoted as

$$E\left(\epsilon_{it}^2\right) = \sigma_t^2,$$

[11] I should formally have a j subscript on wages but since it does not add clarity I have dropped it.

are allowed to change with t and to differ from the conditional variances

$$E\left(\epsilon_{it}^2 | y_i^{t-1}, \eta_i\right).$$

As before, we could consider unobserved heterogeneity components in those conditional variances, both at the individual and job-specific level.

3.3.2 Specification by Type of Exit

In the empirical analysis I will also consider an extended specification that reflects different dynamics across individuals and time according to the type of job change

$$y_{it} = \alpha y_{it-1} + \beta_l d_{it-1}^{loss} y_{it-1} + \beta_q d_{it-1}^{quit} y_{it-1} + \mu_i + \phi_{ij} + \epsilon_{it};\ (i = 1, ...N; t = 2, .., T_i), \quad (3.3)$$

where d_{it}^{loss} is a dummy variable equal to one if worker i at time t ends current job due to an involuntary job separation or job loss; and d_{it}^{quit} equal one if worker i at time t ends current job because she has decided to moved to a new job.

I consider the kind of individual and stochastic effects which preserve the same properties as the basic specification.

3.4 Identification and Estimation method

In this section I discuss the conditions under which I achieve parameter identification. In the model, wages are observed conditional on individuals working; within-job wages, which identifies the parameter α and the individual component η_i, are only observed if the individual does not change job; between-job wage growth, which helps identify heterogeneity across jobs, ϕ_{ij}, and differences on dynamics on years of change, β, is observed only for job movers. Further, participation and mobility decisions can be all endogenous and if this is ignored we risk biasing the estimates of the model[12]. Regarding participation, given the type of individuals considered in the sample, it does not seem such a big issue in this setting so I will ignore it. The potential endogeneity of job mobility is circumvented by controlling for possibly correlated individual and job-specific heterogeneity, without observing it, and by means of a instrument variable estimation method[13].

3.4.1 Orthogonality Conditions

As a matter of notation, I assume that the first observation occurs at $t = 1$, so that the earnings equation (3.1) rewritten in first differences is defined from $t = 3$

$$\Delta y_{it} = \alpha \Delta y_{it-1} + \beta \Delta \left(d_{it-1} y_{it-1} \right) + \left(\eta_{i(t)} - \eta_{i(t-1)} \right) + \Delta \epsilon_{it};\ (i = 1, ...N; t = 3, .., T_i).$$

Given (3.2), the following moment conditions hold

$$E\left(y_i^{t-2} (1 - d_{it-1}) \Delta \epsilon_{it} \right) = 0;\ (t = 3, .., T_i),$$

and so

$$E\left(y_i^{t-2} (1 - d_{it-1}) \left(\Delta y_{it} - \alpha \Delta y_{it-1} - \beta \Delta \left(d_{it-1} y_{it-1} \right) \right) \right) = 0. \tag{3.4}$$

Then, we can consider GMM estimators that used all the available lags at each period as instruments for the equations in first differences (Holtz-Eakin, Newey, and Rosen, 1988;

---

[12] As pointed out by Low *et al.* (2007) this, implicitly, has been the assumption made in papers estimating the covariance structure of earnings (MaCurdy, 1982; Abowd and Card, 1989; Meghir and Pistaferri, 2004).

[13] Low et al. (2007) use a similar sample selection procedure and consider a specification for the wage process fully parametric. Given the distributional assumption, in the estimation they control for selection into employment and for job mobility using the Heckman 2-step method. They claim that: “It is clear that what really matters is the firm mobility decision. Indeed, neglecting the participation correction reduces the variances of interest but the effects are minuscule.”

Arellano and Bond, 1991). Notice that GMM estimation will only consider the moment conditions with $d_{it-1} = 0$, and β would be identified thanks to those with $d_{it-1} = 0$ but $d_{it-2} = 1$.

3.4.2 GMM Estimation

The GMM estimator of $\theta = (\alpha, \beta)'$ based on the corresponding sample moments for (3.4) with weight matrix A_N is given by

$$\hat{\theta}_{GMM} = \arg\min_{\theta} \left[\sum_{i=1}^{N} \Delta v_i' Z_i \right] A_N \left[\sum_{i=1}^{N} Z_i' \Delta v_i \right]$$

where $v_i = y_i - W_i\theta$, with $y_i = (y_{i3}, \ldots, y_{iT_i})'$, $W_i = \begin{pmatrix} y_{i2} & d_{i2}y_{i2} \\ \vdots & \vdots \\ y_{iT_i-1} & d_{iT_i-1}y_{iT_i-1} \end{pmatrix}$, and

$$Z_i = \begin{pmatrix} y_{i1}(1-d_{i2}) & & & 0 \\ & (y_{i1}, y_{i2})(1-d_{i3}) & & \\ & & \ddots & \\ 0 & & & (y_{i1}, \ldots, y_{iT_i-2})(1-d_{iT_i-1}) \end{pmatrix}.$$

According to standard GMM theory an optimal choice of the inverse weight matrix, $V_N = A_N^{-1}$, is a consistent estimate of the covariance matrix of the orthogonality conditions $E(Z_i'\Delta v_i \Delta v_i' Z_i)$. A one-step GMM estimator uses

$$\hat{V} = \sum_{i=1}^{N} Z_i' D D' Z_i$$

where D is the first-difference matrix operator, and a two-step GMM estimator uses the robust choice

$$\tilde{V} = \sum_{i=1}^{N} Z_i' \Delta\hat{v}_i \Delta\hat{v}_i' Z_i,$$

where $\Delta\hat{v}_i$ are one-step residuals.

An estimate of the asymptotic variance of two-step GMM is given by

$$\widehat{Var}\left(\hat{\theta}_{GMM2}\right) = \left[\left(\sum_{i=1}^{N} \Delta W_i' Z_i\right) \tilde{V}^{-1} \left(\sum_{i=1}^{N} Z_i' \Delta W_i\right)\right]^{-1}.$$

3.5 Estimation Results

In this section I show the results corresponding to the GMM estimation of the two specifications presented in Section 3.3 (equations 3.1 and 3.3). In both cases, the dependent variable that I use in the estimation, y_{it}, are log annual real wages residuals from first stage regressions on year dummies, age, education, dummies for race (white), region of residence, and residence in a SMSA[14].

3.5.1 Common Parameters Estimates

I begin by obtaining alternative estimates of a univariate AR(1) model (setting $\beta = 0$). Table C.8 compares OLS in levels, first differences, and within- groups with those obtaining by GMM, using as instruments for the equation in first differences of the lags of wages up to $t-2$. Taking GMM as a benchmark (columns 4 and 5), OLS in levels is biased upward and OLS in differences biased downward, as we would expect for an AR data generating process with individual unobserved heterogeneity. However, the comparison

[14] In earnings dynamics research it is standard to adopt a two step procedure. In the first stage regression, the log of real wages is regressed on control variables and year dummies to eliminate group heterogeneities and aggregate time effects. Then, in the second stage, the unobserved heterogeneity and dynamics of the residuals are modelled.

with the WG is puzzling, since we would also expect a downward bias in that case. Although the system-GMM estimate is bigger than WG, the Sargan test rejects the mean stationarity. Finally, the two-step AR(2) estimates reported in the last column do not change the conclusions, that suggests misspecification as a likely reason for these results[15].

Model in equation (3.1) differs from the previous standard AR(1) model in two main aspects: the different dynamics within and between jobs and the individual-job specific unobserved heterogeneity. The first two columns in Table C.9 report GMM estimates (one- and two-step) of the basic specification, and column 3 corresponds to the two-step GMM estimates of the specification by type of exit. For comparison, I also include GMM estimates for a specification setting $\beta = 0$ (column 4) and another ignoring job-specific heterogeneity (column 5).

Controlling for individual and job-specific effects, GMM estimates of the AR coefficient within groups, α, are almost zero; and across jobs, β, is significant but small (columns 1 and 2). The corresponding estimates for the AR coefficients when I distinguish between involuntary, β_l, and voluntary changes, β_q, are very close to each other (the difference is statistically insignificant). If I impose the same dynamics, both within and between jobs, but still allowing for individual and job-specific effects, the $\hat{\alpha}$ estimate increases capturing the effect of job mobility (column 4). Finally, if I ignore the possibility of heterogeneous match effects across jobs the results for $\hat{\alpha}$ and $\hat{\beta}$ show a marked discrepancy between columns 5 and 2 (my preferred specification).

3.5.2 Variance estimates

The individual-job specific effects for a given individual i and job j, $\eta_{ij} = \mu_i + \phi_{ij}$, can be estimated as

$$\hat{\eta}_{ij} = \frac{1}{T_{ij} - 2} \sum_{s=3}^{T_{ij}} \hat{v}_{is},$$

[15] These results are in line with the ones in Alvarez and Arellano (2004).

where $\hat{v}_{is} = y_{is} - \hat{\alpha} y_{is-1} - \hat{\beta} d_{is-1} y_{is-1}$. They will typically be very noisy since the number of observations for individual-job spells is small. However, the variance of η_{ij} can still be consistently estimated for large N. Optimal estimation of σ_η^2, σ_μ^2 and σ_ϕ^2 requires consideration of the data covariance structure. The errors in levels, $v_{it} = \mu_i + \phi_{ij} + \epsilon_{it}$, satisfy $Var\left(v_{it}\right) = \sigma_\eta^2 + \sigma_t^2$, and $Cov(v_{it}, v_{is}) = \sigma_\eta^2$. Moreover, if we assume that - once we have controlled for μ_i - it would not make much sense to consider correlations across jobs and correlations between μ_i and ϕ_{ij}, the errors would satisfy $Var\left(v_{it}\right) = \sigma_\mu^2 + \sigma_\phi^2 + \sigma_t^2$, and

$$Cov(v_{it}, v_{is}) = \begin{cases} \sigma_\mu^2 + \sigma_\phi^2 & \text{if same job at time } t \neq s, \\ \sigma_\mu^2 & \text{if different job at time } t \neq s. \end{cases}$$

so, simple consistent estimates can be obtained combining sample covariances as

$$\widehat{\left(\sigma_\mu^2 + \sigma_\phi^2\right)} = \sum_{r=1}^{T-2} \left[\frac{1}{T-r-1} \sum_{t=r+2}^{T} \frac{1}{\sum_{i=1}^{N} S_{it}} \sum_{i=1}^{N} S_{it} \hat{v}_{it} \hat{v}_{it-r} \right],$$

and

$$\widehat{\sigma}_\mu^2 = \sum_{r=1}^{T-2} \left[\frac{1}{T-r-1} \sum_{t=r+2}^{T} \frac{1}{\sum_{i=1}^{N} (1 - S_{it})} \sum_{i=1}^{N} \left(1 - S_{it}\right) \hat{v}_{it} \hat{v}_{it-r} \right],$$

where $S_{it} = \prod_{s=1}^{r} (1 - d_{it-s}) = (1 - d_{it-1}) \cdot (1 - d_{it-2}) \cdot \ldots \cdot (1 - d_{it-r})$ indicates that individual i stays at the same job between $t - r$ and t.

The results are reported in Table C.10. I find that in the whole sample (column 1) the estimated variance of the individual effects is 0.09, very close to the variance of the sum of these and the job-specific effects, because for the stayers (people who never change job) it is not possible to discriminate between those two components (column 2). If I only consider individuas that change at least once (column 3), the estimated variance of the job-specific effects represents around one third of the variance for the individual fixed effects. Finally, if I only use those who suffer involuntary job changes (column 4) the variance of the heterogeneity across jobs increases up to one half.

3.6 Conclusions

The chapter develops an error components model designed to more thoroughly describe the impact of job mobility on the dynamics and heterogeneity of individual wages than previous references. In particular, the specification proposed has two different parameters to capture dynamics within jobs and across jobs, and the unobserved heterogeneity shows a richer pattern, as well, composed of both individual and job-specific effects. The potential endogeneity of job mobility in relation to earnings is circumvented using a Generalized Method of Moments estimation that controls for those two unobserved heterogeneity components.

In the data, drawn from the PSID, I find that - once we control for individual and job-specific effects - the dynamics within jobs is almost zero, whereas across jobs is significant but small. For the dynamics, the distinction between voluntary and involuntary transitions turns out to be irrelevant. However, that distinction matters in the case of the components of the cross-sectional variance. The estimated variance of the job-specific effects represents around one third of the variance for the individual fixed effects. If I consider a subsample that only includes involuntary job changes, the estimated variance of the heterogeneity across jobs increases up to one half.

Further research is needed on the consideration in the model of the labour market participation decision and, thus, the inclusion of women and transitions job-to-nonemployment and nonemployment-to-job into the analysis.

Appendix of Chapter 3

3.A Sample Selection

Starting point: PSID 1968-1993 Family and Individual - merged files (53,005 individuals).

1. Drop members of the Latino sample (10,022 individuals) = Sample (42,983 individuals).

2. Keep only those who are continuously heads of their households = Sample (16,038 individuals).

3. Keep only males aged 25 to 55 over the period = Sample (8,190 individuals).

4. Drop those with a spell of self-employment = Sample (6,303 individuals).

5. Drop those with missing race, education and region of residence records = Sample (6,047 individuals).

6. Drop those with top-coded earnings records and those with missing earnings = Sample (5,479 individuals).

7. Drop those with outlying earnings records, that is, a change in log earnings greater than 5 or less than -3 = Sample (5,384 individuals).

8. Drop those with missing records on reason of job change question and those with noncontinuous data = Sample (5,345 individuals).

9. Keep only those who are in the sample for 8 years or more

= FINAL SAMPLE: Males, 1968-1992 (2,013 individuals and 27,845 observations).

Chapter 4

Estimating Nonlinear Models with Multiple Fixed Effects: A Computational Note[1]

4.1 Introduction

In a typical nonlinear micropanel data model with fixed effects there are hundreds or thousands of individual coefficients to estimate together with a relatively small number of common parameters. A well known computational simplification in the linear model is to obtain first the maximum likelihood (ML) estimates of the common parameters from a regression on the data in deviations from individual means, and secondly retrieve ML estimates of the effects from averaged residuals one by one. A similar computational simplification is available for Newton-Raphson and related algorithms for nonlinear fixed effects models, which exploits the block-diagonal structure of the Hessian. This simplification has been discussed in Hall (1978), Chamberlain (1980), and Greene (2004) for nonlinear models with a scalar fixed effect. The first purpose of this work is to show how to use an iterated algorithm of this type in a nonlinear model with multiple fixed effects.

As first noted by Neyman and Scott (1948), when the time series dimension T is small

[1] This chapter is part of a joint work with Manuel Arellano.

relative to the cross-sectional dimension n, ML estimates of the common parameters can be severely biased, specially in dynamic models. This Incidental Parameters problem arises because the unobserved individual characteristics are replaced by noisy estimates, which bias estimates of model parameters. In particular, the bias of the MLE is of order $1/T$. In some special cases it is possible to obtain fixed T large n consistent estimators of certain common parameters, but these situations are more the exception than the rule. Alternatively, a number of additional approaches have been proposed to obtain approximately unbiased estimators as opposed to estimators with no bias at all[2]. One of these approaches consists of estimation from a bias corrected objective function relative to some target criterion[3]. In this chapter we discuss the application of computationally efficient algorithms to modified concentrated likelihoods of this type to obtain estimators without bias to order $1/T$ in nonlinear panel models with multiple fixed effects.

The chapter is organized as follows. Section 4.2 introduces the model and notation. Section 4.3 explains how the iterated algorithm works. Section 4.4 discusses its application to bias corrected concentrated likelihoods. Section 4.5 presents some simulation results. Finally, Section 4.6 concludes. Detailed derivations are given in the Appendix.

4.2 Model and Notation

Let us consider the following model for the joint density of T random vectors conditioned on initial observations, strictly exogenous variables, and fixed effects:

$$f\left(y_{i1}, ..., y_{iT} \mid y_{i0}, x_{i1}, ..., x_{iT}, \alpha_{i0}\right) = \prod_{t=1}^{T} f\left(y_{it} \mid y_{i(t-1)}, x_{it}, \alpha_{i0}, \theta_0\right)$$

---

[2]See Arellano and Hahn, 2006a, for a review of this literature on bias-adjusted estimation methods for nonlinear panel data models with fixed effects.

[3]See Pace and Salvan (2005) for adjustments of this type for a generic concentrated likelihood with independent observations, Arellano and Hahn (2006a) for static nonlinear panel models and Arellano and Hahn (2006b) and the second chapter of this thesis for the dynamic case.

where θ_0 is a vector of common parameters and α_{i0} is a vector of fixed effects. We observe the random sample $\{y_{i0}, ..., y_{iT}, x_{i0}, ..., x_{iT}\}_{i=1}^{n}$ and we denote $\alpha_0 = (\alpha_{10}', ..., \alpha_{n0}')'$ and $\delta_0 = (\theta_0', \alpha_0')'$. Let the log likelihood of one observation be

$$\ell_{it}(\theta, \alpha_i) = \ln f\left(y_{it} \mid y_{i(t-1)}, x_{it}, \alpha_i, \theta\right)$$

and let $\ell_i(\theta, \alpha_i) = \sum_{t=1}^{T} \ell_{it}(\theta, \alpha_i)$.

4.3 Efficient Newton-Raphson iteration

Let us consider the estimator

$$\begin{pmatrix} \widehat{\theta} \\ \\ \widehat{\alpha} \end{pmatrix} = \arg\max_{\theta,\alpha} \sum_{i=1}^{n} \ell_i(\theta, \alpha_i)$$

and let first and second derivatives be denoted by

$$\begin{aligned} d_{\theta i} &= \frac{\partial \ell_i(\theta, \alpha_i)}{\partial \theta}, \quad d_{\alpha i} = \frac{\partial \ell_i(\theta, \alpha_i)}{\partial \alpha_i} \\ H_{\theta\theta i} &= \frac{\partial^2 \ell_i(\theta, \alpha_i)}{\partial \theta \partial \theta'}, \quad H_{\alpha\alpha i} = \frac{\partial^2 \ell_i(\theta, \alpha_i)}{\partial \alpha_i \partial \alpha_i'}, \quad H_{\theta\alpha i} = \frac{\partial^2 \ell_i(\theta, \alpha_i)}{\partial \theta \partial \alpha_i'}. \end{aligned}$$

The Kth step of the iteration of a computationally efficient algorithm for obtaining $\widehat{\theta}$ and $\widehat{\alpha}$ takes the form

$$\theta_{[K]} - \theta_{[K-1]} = -\left[\sum_{i=1}^{n}\left(H_{\theta\theta i} - H_{\theta\alpha i} H_{\alpha\alpha i}^{-1} H_{\alpha\theta i}\right)\right]^{-1} \sum_{i=1}^{n}\left(d_{\theta i} - H_{\theta\alpha i} H_{\alpha\alpha i}^{-1} d_{\alpha i}\right) \tag{4.1}$$

$$\alpha_{i[K]} - \alpha_{i[K-1]} = -H_{\alpha\alpha i}^{-1}\left[d_{\alpha i} + H_{\alpha\theta i}\left(\theta_{[K]} - \theta_{[K-1]}\right)\right], \quad (i = 1, ..., n) \tag{4.2}$$

where all derivatives are evaluated at $\theta_{[K-1]}$ and $\alpha_{i[K-1]}$.

This result can be easily proved using partitioned inverse formulae (a detailed derivation is in the Appendix 4.A). It is a standard result in nonlinear estimation of models with many group effects.[4]

4.4 Adjusted Concentrated Likelihood

When T is short we may be interested to consider an estimator that maximizes a bias corrected concentrated likelihood of the type reviewed in Arellano and Hahn (2006a):

$$\widehat{\theta}_c = \arg\max_{\theta} \sum_{i=1}^{n} \left[\ell_i\left(\theta, \widehat{\alpha}_i\left(\theta\right)\right) + \beta_i\left(\theta, \widehat{\alpha}_i\left(\theta\right)\right)\right]$$

where

$$\widehat{\alpha}_i\left(\theta\right) = \arg\max_{\alpha} \ell_i\left(\theta, \alpha\right)$$

and $\beta_i\left(\theta, \alpha_i\right)$ is an adjustment term.

As long as the adjustment term depends on α, the iterated algorithm discussed above cannot be directly used for estimating $\widehat{\theta}_c$. Note that

$$\begin{pmatrix} \widehat{\theta}_c \\ \widehat{\alpha}_c \end{pmatrix} = \arg\max_{\theta,\alpha} \sum_{i=1}^{n} \left[\ell_i\left(\theta, \alpha_i\right) + \beta_i\left(\theta, \widehat{\alpha}_i\left(\theta\right)\right)\right]$$

where $\widehat{\alpha}_c = \widehat{\alpha}\left(\widehat{\theta}_c\right)$. Thus, if we use the analysis of covariance algorithm discussed in the previous section we still need to calculate $\widehat{\alpha}_i\left(\theta\right)$ for given values of θ.

[4] An alternative Gauss-Newton algorithm which leads to a regression-based iteration is discussed in Appendix 4.B.

An Alternative, Computationally Effective Estimator Alternatively, we can consider an estimator of the form

$$\begin{pmatrix} \widetilde{\theta} \\ \widetilde{\alpha} \end{pmatrix} = \arg\max_{\theta,\alpha} \sum_{i=1}^{n} \left[\ell_i\left(\theta, \alpha_i\right) + \beta_i\left(\theta, \alpha_i\right)\right]$$

for which the iterated algorithm can be used. This is equivalent to:

$$\widetilde{\theta} = \arg\max_{\theta} \sum_{i=1}^{n} \left[\ell_i\left(\theta, \widetilde{\alpha}_i\left(\theta\right)\right) + \beta_i\left(\theta, \widetilde{\alpha}_i\left(\theta\right)\right)\right]$$

where

$$\widetilde{\alpha}_i\left(\theta\right) = \arg\max_{\alpha} \left[\ell_i\left(\theta, \alpha\right) + \beta_i\left(\theta, \alpha_i\right)\right].$$

The statistic $\widetilde{\alpha}_i\left(\theta\right)$ can be regarded as a Bayesian estimator that uses $e^{\beta_i(\theta,\alpha_i)}$ as the prior distribution of α_i for a given value of θ. Thus, under general conditions, $\widetilde{\alpha}_i\left(\theta\right)$ will be asymptotically equivalent to $\widehat{\alpha}_i\left(\theta\right)$, and $\widetilde{\theta}$ will have similar (bias reducing) properties as $\widehat{\theta}$ (see Severini, 1998, section 4, for a discussion on the use of adjusted concentrated likelihoods using alternative estimates of nuisance parameters).

It appears that $\widetilde{\theta}$ is not only computationally convenient, but it may also exhibit improved finite sample properties in certain situations due to the replacement of $\widehat{\alpha}_i\left(\theta\right)$ by $\widetilde{\alpha}_i\left(\theta\right)$.

4.5 Monte Carlo Study

In this section Monte Carlo simulations are used to evaluate the performance of the efficient algorithm in different sample sizes, and its application to bias corrected concentrated likelihoods of nonlinear models. We consider four examples in this section, but keeping the simulation design as consistent as possible across the models: the static probit with scalar fixed effects, the dynamic probit with scalar fixed effects, the static probit with multiple fixed effects and the dynamic probit with multiple fixed effects. Thus,

$$y_{it} = \mathbf{1}\left[w_{it} + \epsilon_{it} > 0\right]$$

where $\epsilon_{it} \sim N(0,1)$. We compare estimates of common parameters estimated by ML and bias-corrected ML[5].

4.5.1 Probit designs with Scalar fixed effects

We consider five different data-generating processes: three for a static probit and two more for a dynamic probit.

Static Probit

- Design 1 (Bester and Hansen, 2005):

$$w_{it}^{(1)} = \eta_{i0} + \beta_0 x_{it}$$

with $\eta_{i0} \sim N(x_{i0}, 1)$, $x_{it} = \frac{1}{2}x_{it-1} + u_{it}$, $u_{it} \sim N(0,1)$, $x_{i0} \sim N(0,1)$, and $\beta_0 = 1$. Models were fit with $T = \{8, 12\}$ and $N = 100$. Each model was fit 1,000 times with random draws for ϵ_{it}. The conditioning data, x_{it}, and η_{i0} were held constant.

- Design 2 (Greene, 2004):

$$w_{it}^{(2)} = \eta_{i0} + \beta_0 x_{it} + \delta_0 d_{it}$$

[5] Other studies, that consider nonlinear designs with scalar fixed effects (Bester and Hansen, 2005; Carro, 2006; and Fernández-Val, 2005), show that the bias in the ML estimator is similar in magnitude for the logit and the probit models and that bias corrections also perform similarly. Here, we focus on probit designs and extend the analysis to consider multiple fixed effects.

with $\eta_{i0} = \sqrt{T}\bar{x}_i + a_i$, $a_i \sim N(0,1)$, $x_{it} \sim N(0,1)$, $h_{it} \sim N(0,1)$, $d_{it} = \mathbf{1}[x_{it} + h_{it} > 0]$, $\beta_0 = 1$ and $\delta_0 = \{1, 0.5\}$.

- Design 3:

$$w_{it}^{(3)} = \eta_{i0} + \beta_0 x_{it} + \delta_0 d_{it}$$

with $\eta_{i0} = 0$, $\forall i$, $x_{it} \sim N(0,1)$, $h_{it} \sim N(0,1)$, $d_{it} = \mathbf{1}[x_{it} + h_{it} > 0]$, $\beta_0 = 1$ and $\delta_0 = 0.5$.

For designs 2 and 3 models were fit with $T = \{6, 8, 10, 12\}$ and $N = 1,000$. Each model was fit 100 times with random draws for ϵ_{it}. The conditioning data, x_{it}, d_{it} and η_{i0} were held constant.

Adjusted Concentrated Likelihood We have that

$$\Pr(y_{it} = 1 | w_{it}) = \Phi(w_{it}) = \Phi_{it}$$

where Φ is the normal cdf. For design 1

$$\Pr(y_{it} = 1 | \eta_i, x_i) = \Phi(\eta_i + \beta x_{it}),$$

whereas for designs 2 and 3

$$\Pr(y_{it} = 1 | \eta_i, x_i, d_i) = \Phi(\eta_i + \beta x_{it} + \delta d_{it}).$$

Let's consider $\alpha_i = \eta_i$ and $\theta = \beta$ for design 1, and $\theta = (\beta, \delta)$ for designs 2 and 3. Let the log-likelihood of one observation be

$$\ell_{it}(\theta, \alpha_i) = y_{it} \ln \Phi_{it} + (1 - y_{it}) \ln(1 - \Phi_{it}),$$

and let $\ell_i(\theta,\alpha_i) = \sum_{t=1}^{T} \ell_{it}(\theta,\alpha_i)$. We can obtain the MLE, $\widehat{\theta}$, as the argument that maximizes the log-likelihood function

$$\begin{pmatrix} \widehat{\theta} \\ \widehat{\alpha} \end{pmatrix} = \arg\max_{\theta,\alpha} \sum_{i=1}^{n} \ell_i(\theta,\alpha_i).$$

Or, equivalently, $\widehat{\theta}$, is the estimator that maximizes the concentrated log-likelihood function:

$$\widehat{\theta} = \arg\max_{\theta} \sum_{i=1}^{n} \left[\ell_i(\theta, \widehat{\alpha}_i(\theta))\right]$$

where

$$\widehat{\alpha}_i(\theta) = \arg\max_{\alpha} \ell_i(\theta,\alpha).$$

The corrected concentrated MLE, $\widehat{\theta}_c$, is the argument that maximizes

$$\widehat{\theta}_c = \arg\max_{\theta} \sum_{i=1}^{n} \left[\ell_i(\theta, \widehat{\alpha}_i(\theta)) + \beta_i(\theta, \widehat{\alpha}_i(\theta))\right]$$

where $\beta_i(\theta,\alpha_i)$ is an adjustment term. The corrected computationally efficient MLE, $\widetilde{\theta}$, is an estimator of the form

$$\begin{pmatrix} \widetilde{\theta} \\ \widetilde{\alpha} \end{pmatrix} = \arg\max_{\theta,\alpha} \sum_{i=1}^{n} \left[\ell_i(\theta,\alpha_i) + \beta_i(\theta,\alpha_i)\right]$$

or equivalently,

$$\widetilde{\theta} = \arg\max_{\theta} \sum_{i=1}^{n} \left[\ell_i(\theta, \widetilde{\alpha}_i(\theta)) + \beta_i(\theta, \widetilde{\alpha}_i(\theta))\right]$$

where

$$\widetilde{\alpha}_i(\theta) = \arg\max_{\alpha} \left[\ell_i(\theta,\alpha) + \beta_i(\theta,\alpha_i)\right].$$

Following Arellano and Hahn (2006a), for a static model with scalar fixed effects, the form of the adjustment term will be:

$$\beta_i(\theta, \alpha_i) = -\frac{1}{2}\left(-\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ell_{it}(\theta, \alpha_i)}{\partial \alpha_i^2}\right)^{-1}\frac{1}{T}\sum_{t=1}^{T}\left[\frac{\partial \ell_{it}(\theta, \alpha_i)}{\partial \alpha_i}\right]^2.$$

Table D.1 lists the means of the empirical sampling distribution for the ML and bias-corrected ML estimators (BC-C for the corrected concentrated MLE, $\widehat{\theta}_c$, and BC-E for the corrected computationally efficient MLE, $\widetilde{\theta}$) in Design 1. The most relevant feature is the upward bias in the ML estimates of β. For each choice of T, the bias corrected estimates perform better both in terms of the bias and the precision. The results for the ML and the BC-C are consistent with the ones in Bester and Hansen (2005). Moreover, the BC-E is slightly better, in addition to the improvement in terms of computational time.

Tables D.2 and D.3 display the means of the empirical sampling distribution for the ML and bias-corrected ML estimators in Design 2 with $\delta_0 = 1$ and $\delta_0 = 0.5$, respectively. This design was proposed by Greene (2004) in order to examine the small sample bias of the fixed effects MLE. For all values of T, the BC estimates offer substantial improvements over ML (again, both in the bias and in the SD). As we would expect, all the estimates improve quickly with the number of time periods. Also we observe that size distortions are bigger for $\delta_0 = 1$ than for $\delta_0 = 0.5$, but corrections perform well in any case. What it is important is that we obtain considerable bias reduction for a T as small as 6 or 8.

Table D.4 lists the means of the empirical sampling distribution for the ML and bias-corrected ML estimators in Design 3. In this design, we simulate the data generating process with $\eta_{i0} = 0, \forall i$, but we estimate as if we would have individual heterogeneity in the data. Even in this case the fixed effects MLE is severely biased. BC estimates offer a great improvement over ML.

Dynamic Probit

- Design 4 (Bester and Hansen, 2005):

$$w_{it}^{(4)} = \eta_{i0} + \beta_0 x_{it} + \delta_0 y_{it-1}$$

with $\eta_{i0} \sim N(x_{i0}, 1)$, $x_{it} = \frac{1}{2}x_{it-1} + u_{it}$, $u_{it} \sim N(0,1)$, $x_{i0} \sim N(0,1)$, $y_{i0} = \mathbf{1}[\eta_{i0} + \beta_0 x_{i0} + \epsilon_{i0} > 0]$, $\beta_0 = 1$, and $\delta_0 = 0.5$. Models were fit with $T = \{8, 12\}$ and $N = 100$. Each model was fit 200 times with random draws for ϵ_{it}. The conditioning data, x_{it}, and η_{i0} were held constant.

- Design 5:

$$w_{it}^{(5)} = \eta_{i0} + \beta_0 x_{it} + \delta_0 y_{it-1}$$

with $\eta_{i0} = 0$, $\forall i$, $x_{it} \sim N(0,1)$, $y_{i0} = \mathbf{1}[\eta_i + \beta_0 x_{i0} + \epsilon_{i0} > 0]$, $\beta_0 = 1$, and $\delta_0 = 0.5$.

Models were fit with $T = \{6, 8, 10, 12\}$ and $N = 1,000$. Each model was fit 100 times with random draws for ϵ_{it}. The conditioning data, x_{it}, and η_{i0} were held constant.

Adjusted Concentrated Likelihood Now, we have that

$$\Pr(y_{it} = 1 | y_{it-1}, \eta_i, x_i) = \Phi(\alpha_i + \beta x_{it} + \delta y_{it-1}).$$

Let's consider $\theta = (\beta, \delta)$ and $\alpha_i = \eta_i$. In the dynamic case, following Arellano and Hahn (2006b) we will consider two different adjustment terms: $\beta_{1i}(\theta, \alpha_i)$ and $\beta_{2i}(\theta, \alpha_i)$.

1. Trimming: In this approach the form of the adjustment term is

$$\beta_{1i}(\theta, \alpha_i) = -\frac{1}{2}(H_i(\theta, \alpha_i))^{-1} \Upsilon_i(\theta, \alpha_i)$$

where

$$H_i(\theta, \alpha_i) = -\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ell_{it}(\theta, \alpha_i)}{\partial \alpha_i^2}$$

and $\Upsilon_i(\theta, \alpha_i)$ is a trimmed version of the outer product term, that is,

$$\Upsilon_i(\theta, \alpha_i) = \Omega_0 + 2\sum_{l=1}^{r}\Omega_l,$$

where

$$\Omega_l = \frac{1}{T-l}\sum_{t=l+1}^{T}\left(1 - \frac{l}{r+1}\right)\frac{\partial \ell_{it}}{\partial \alpha_i}\cdot\frac{\partial \ell_{it-l}}{\partial \alpha_i},$$

for r small.

2. Expected Quantities: This approach is based in a local version of the estimated bias using that at the truth $H_i^{-1}(\Gamma_0)\Upsilon_i(\Gamma_0) = 1$ (Pace and Salvan, 2005)[6]. Now the form of the adjustment term is

$$\beta_{2i}(\theta, \alpha_i) = \frac{1}{2}\ln(H_i(\theta, \alpha_i)) - \frac{1}{2}\ln(\Upsilon_i(\theta, \alpha_i))$$

where, as before,

$$H_i(\theta, \alpha_i) = -\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2 \ell_{it}(\theta, \alpha_i)}{\partial \alpha_i^2}$$

but now $\Upsilon_i(\theta, \alpha_i)$ is based on the expectation

$$\bar{\Upsilon}_i(\theta, \alpha_i; \theta_0, \alpha_{i0}) = TE_0\left\{\left[\frac{\partial \ell_i(\alpha, \Theta_i)}{\partial \alpha_i} - E\left(\frac{\partial \ell_i(\alpha, \Theta_i)}{\partial \alpha_i}\right)\right]^2 \middle| y_{i0}\right\}.$$

What we calculate in practice is $\Upsilon_i(\theta, \alpha_i) = \bar{\Upsilon}_i\left(\theta, \alpha_i; \hat{\theta}, \widehat{\alpha}_i\right)$, because true values, θ_0 and α_{i0}, are unknow. We obtain $\bar{\Upsilon}_i\left(\theta, \alpha_i; \hat{\theta}, \widehat{\alpha}_i\right)$ as a mean of $\{\Upsilon_i^m(\theta, \alpha_i)\}_{m=1}^{M}$

[6] This local version thus is appropriate in a likelihood context.

calculated in M samples simulated as $\left\{\prod_{t=1}^{T} f\left(y_{it}|y_{i0},\hat{\theta},\widehat{\alpha}_i\right)\right\}_{i=1}^{N}$. That is,

$$\bar{\Upsilon}_i\left(\theta,\alpha_i;\hat{\theta},\widehat{\alpha}_i\right)=\frac{1}{M}\sum_{m=1}^{M}\Upsilon_i^m\left(\theta,\alpha_i\right),$$

where

$$\Upsilon_i^m\left(\theta,\alpha_i\right)=\frac{1}{T}\sum_{t=1}^{T}\left[\frac{\partial\ell_{it}}{\partial\alpha_i}-\left(\frac{1}{T}\sum_{r=1}^{T}\frac{\partial\ell_{ir}}{\partial\alpha_i}\right)\right]^2.$$

Summing up, BC-C-Trimming estimator, $\widehat{\theta}_{1c}$, and BC-C-Expectation estimator, $\widehat{\theta}_{2c}$, maximize

$$\widehat{\theta}_{jc}=\arg\max_{\theta}\sum_{i=1}^{n}\left[\ell_i\left(\theta,\widehat{\alpha}_i\left(\theta\right)\right)+\beta_{ji}\left(\theta,\widehat{\alpha}_i\left(\theta\right)\right)\right],\ \text{with}\ \ (j=1,2)\,.$$

And BC-E-Trimming, $\widetilde{\theta}_1$, BC-E-Expectation, $\widetilde{\theta}_2$, are estimators of the form

$$\begin{pmatrix}\widetilde{\theta}_j\\ \widetilde{\alpha}_j\end{pmatrix}=\arg\max_{\theta,\alpha}\sum_{i=1}^{n}\left[\ell_i\left(\theta,\alpha_i\right)+\beta_{ji}\left(\theta,\alpha_i\right)\right],\ \text{with}\ \ (j=1,2)$$

or equivalently,

$$\widetilde{\theta}_j=\arg\max_{\theta}\sum_{i=1}^{n}\left[\ell_i\left(\theta,\widetilde{\alpha}_i\left(\theta\right)\right)+\beta_{ji}\left(\theta,\widetilde{\alpha}_i\left(\theta\right)\right)\right],\ \text{with}\ \ (j=1,2)\,.$$

Table D.5 lists the means of the empirical sampling distribution for the ML and bias-corrected ML estimators in Design 4. Our results for Design 4 are comparable with the ones in Bester and Hansen (2005). We contribute by adding the corrections based in expected terms (both the concentrated and the computationally effective versions), which are the best ones among the proposed estimators.

Table D.6 lists the means of the empirical sampling distribution for the ML and bias-corrected ML estimators in Design 5. Although the biases are smaller with respect to design 4, bias corrected estimates perform better than ML. The correction based on expected terms displays by far the smallest size distortions, and this is specially true for the BC-E-Exp.

4.5.2 Probit designs with Multiple fixed effects

In this section we consider two data-generating processes more, one static probit and another dynamic, but with multiple individual fixed effects.

Static Probit

- Design 6:

$$w_{it}^{(6)} = \eta_{i0} + \beta_0 x_{it} + \delta_{i0} d_{it}$$

  where $x_{it} \sim N(0,1)$, $d_{it} = \mathbf{1}[x_{it} + h_{it} > 0]$, $h_{it} \sim N(0,1)$, $\eta_{i0} = 0$ and $\delta_{i0} = 0.5$ $\forall i$, and $\beta_0 = 1$.

Models were fit with $T = \{6, 8, 10, 12, 20\}$ and $N = 1,000$. Each model was fit 100 times with random draws for ϵ_{it}. The conditioning data, x_{it}, d_{it}, η_{i0}, and δ_{i0} were held constant.

Adjusted Concentrated Likelihood Now, we have that

$$\Pr(y_{it} = 1 | \eta_i, \delta_i, x_i, d_i) = \Phi(\eta_i + \beta x_{it} + \delta_i d_{it}).$$

Let's consider $\theta = \beta$ and $\alpha_i = (\eta_i, \delta_i)$. As before, the BC-C, $\widehat{\theta}_c$, is the one that maximizes

$$\widehat{\theta}_c = \arg\max_{\theta} \sum_{i=1}^{n} \left[\ell_i(\theta, \widehat{\alpha}_i(\theta)) + \beta_i(\theta, \widehat{\alpha}_i(\theta))\right]$$

and the BC-E, $\widetilde{\theta}$, is an estimator of the form

$$\begin{pmatrix} \widetilde{\theta} \\ \widetilde{\alpha} \end{pmatrix} = \arg\max_{\theta,\alpha} \sum_{i=1}^{n} \left[\ell_i\left(\theta, \alpha_i\right) + \beta_i\left(\theta, \alpha_i\right) \right]$$

or equivalently,

$$\widetilde{\theta} = \arg\max_{\theta} \sum_{i=1}^{n} \left[\ell_i\left(\theta, \widetilde{\alpha}_i\left(\theta\right)\right) + \beta_i\left(\theta, \widetilde{\alpha}_i\left(\theta\right)\right) \right]$$

where now, for the case of a static model but with multiple fixed effects, the form of the adjustment term $\beta_i\left(\theta, \alpha_i\right)$ is:

$$\beta_i\left(\theta, \alpha_i\right) = -\frac{1}{2} tr \left\{ \left(-\frac{1}{T} \sum_{t=1}^{T} \frac{\partial^2 \ell_{it}\left(\theta, \alpha_i\right)}{\partial \alpha_i \partial \alpha_i'} \right)^{-1} \frac{1}{T} \sum_{t=1}^{T} \left[\frac{\partial \ell_{it}\left(\theta, \alpha_i\right)}{\partial \alpha_i} \cdot \frac{\partial \ell_{it}\left(\theta, \alpha_i\right)}{\partial \alpha_i'} \right] \right\}.$$

Table D.7 lists the means of the empirical sampling distribution for the ML and bias-corrected ML estimators in Design 6. If we compare this table with Table D.4, we can see how adding individual heterogeneity in a robust way into our framework increases the bias in the estimates of β. Now, the BC-E estimate is less biased than the ML and even a bit less than BC-C, specially for smaller T's.

Dynamic Probit

- Design 7:

$$w_{it}^{(7)} = \eta_i + \beta x_{it} + \delta_i y_{it-1}$$

where $x_{it} \sim N\left(0,1\right)$, $y_{i0} = \mathbf{1}[\eta_{i0} + \beta x_{i0} + \epsilon_{i0} > 0]$, $\eta_{i0} = 0$ and $\delta_{i0} = 0.5\ \forall i$, and $\beta_0 = 1$.

Models were fit with $T = \{6, 8, 10, 12, 20\}$ and $N = 1,000$. Each model was fit 100 times with random draws for ϵ_{it}. The conditioning data, x_{it}, η_{i0}, and δ_{i0} were held constant.

Adjusted Concentrated Likelihood Now, we have that

$$\Pr\left(y_{it}=1|y_{it-1},\eta_i,\delta_i,x_i\right)=\Phi\left(\eta_i+\beta x_{it}+\delta_i y_{it-1}\right).$$

Let's consider $\theta=\beta$ and $\alpha_i=(\eta_i,\delta_i)$. With multiple fixed effects the two adjustment terms are the following.

1. Trace Based Approach with Trimming:

$$\beta_{1i}\left(\theta,\alpha_i\right)=-\frac{1}{2}tr\left\{\left(H_i\left(\theta,\alpha_i\right)\right)^{-1}\Upsilon_i\left(\theta,\alpha_i\right)\right\}$$

where

$$H_i\left(\theta,\alpha_i\right)=-\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2\ell_{it}\left(\theta,\alpha_i\right)}{\partial\alpha_i\partial\alpha_i'}$$

and $\Upsilon_i\left(\theta,\alpha_i\right)$ is a trimmed version of the outer product term, that is,

$$\Upsilon_i\left(\theta,\alpha_i\right)=\Omega_0+\sum_{l=1}^{r}\left(\Omega_l+\Omega_l'\right),$$

where

$$\Omega_l=\frac{1}{T-l}\sum_{t=l+1}^{T}\left(1-\frac{l}{r+1}\right)\frac{\partial\ell_{it}}{\partial\alpha_i}\cdot\frac{\partial\ell_{it-l}}{\partial\alpha_i'},$$

for r small.

2. Determinant Based Approach Using Expected Quantities:

$$\beta_{2i}\left(\theta,\alpha_i\right)=\frac{1}{2}\ln\det\left(H_i\left(\theta,\alpha_i\right)\right)-\frac{1}{2}\ln\det\left(\Upsilon_i\left(\theta,\alpha_i\right)\right)$$

where, as before,

$$H_i\left(\theta,\alpha_i\right)=-\frac{1}{T}\sum_{t=1}^{T}\frac{\partial^2\ell_{it}\left(\theta,\alpha_i\right)}{\partial\alpha_i\partial\alpha_i'}$$

and now $\Upsilon_i(\theta, \alpha_i)$ is obtained as

$$\bar{\Upsilon}_i\left(\theta, \alpha_i; \hat{\theta}, \widehat{\alpha}_i\right) = \frac{1}{M}\sum_{m=1}^{M} \Upsilon_i^m(\theta, \alpha_i),$$

where

$$\Upsilon_i^m(\theta, \alpha_i) = \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\left[\frac{\partial \ell_{it}}{\partial \alpha_i} - \left(\frac{1}{T}\sum_{r=1}^{T}\frac{\partial \ell_{ir}}{\partial \alpha_i}\right)\right]\cdot\left[\frac{\partial \ell_{it}}{\partial \alpha_i'} - \left(\frac{1}{T}\sum_{r=1}^{T}\frac{\partial \ell_{ir}}{\partial \alpha_i'}\right)\right].$$

Again, BC-C-Trimming, $\widehat{\theta}_{1c}$, and BC-C-Expectation, $\widehat{\theta}_{2c}$, maximize

$$\widehat{\theta}_{jc} = \arg\max_{\theta}\sum_{i=1}^{n}\left[\ell_i(\theta, \widehat{\alpha}_i(\theta)) + \beta_{ji}(\theta, \widehat{\alpha}_i(\theta))\right], \text{ with } (j = 1, 2).$$

And BC-E-Trimming, $\widetilde{\theta}_1$, and BC-E-Expectation, $\widetilde{\theta}_2$, would be estimators of the form

$$\begin{pmatrix} \widetilde{\theta}_j \\ \widetilde{\alpha}_j \end{pmatrix} = \arg\max_{\theta,\alpha}\sum_{i=1}^{n}\left[\ell_i(\theta, \alpha_i) + \beta_{ji}(\theta, \alpha_i)\right]$$

or equivalently,

$$\widetilde{\theta}_j = \arg\max_{\theta}\sum_{i=1}^{n}\left[\ell_i(\theta, \widetilde{\alpha}_i(\theta)) + \beta_{ji}(\theta, \widetilde{\alpha}_i(\theta))\right], \text{ with } (j = 1, 2).$$

Table D.8 lists the means of the empirical sampling distribution for the ML and bias-corrected ML estimators in Design 7. Adding more fixed effects clearly increases complexity since the model is more demanding in terms of T. This is reflected in the fact that size distortions are bigger than for the scalar case (Table D.6). In any case, we obtain improved estimates applying the corrections based on trimming (again slightly better for the BC-E than BC-C). Regarding the correction based on expected terms, further research is needed in terms of the number of simulated samples required to obtain the expectations since the results are very sensitive to the value of this parameter. In the scalar case with values of this parameter M around 200-300 was enough for obtaining negligible differences across designs.

4.6 Conclusions

In this chapter, we consider estimation of nonlinear panel data models that include multiple individual fixed effects. Estimation of these models is complicated both by the difficulty of estimating models with possibly thousands of coefficients and also by the incidental parameters problem; that is, noisy estimates of the fixed effects when the time dimension is short contaminates the estimates of the common parameters due to the nonlinearity of the problem. We show how to use an iterated algorithm which simplifies estimation in a nonlinear model with multiple fixed effects and we also discuss its application to bias corrected concentrated likelihoods.

Simulations show that the estimator proposed is not only computationally convenient but it is also as good as others in a variety of probit designs. Different adjustments of the likelihood function result in bias corrected estimators that perform comparably to other bias corrections proposed in the literature. We can think in many microeconometric applications that use nonlinear panel data models. The results of the chapter suggest that bias corrected estimates will be very useful in relevant empirical settings given the sample sizes of the panels more often used by researchers and, moreover, because they allow us to introduce more individual heterogeneity to address endogeneity concerns in a robust way.

Appendix of Chapter 4

4.A Newton-Raphson iteration

The Kth step of the Newton-Raphson iteration takes the form

$$\delta_K = \delta_{K-1} - \left(\frac{\partial^2 L\left(\delta_{K-1}\right)}{\partial\delta\partial\delta'}\right)^{-1} \frac{\partial L\left(\delta_{K-1}\right)}{\partial\delta},$$

or for shortness

$$\Delta\delta = -\left(\frac{\partial^2 L}{\partial\delta\partial\delta'}\right)^{-1} \frac{\partial L}{\partial\delta}$$

where $L\left(\delta\right) = \sum_{i=1}^{n} \ell_i\left(\theta, \alpha_i\right)$ and

$$\frac{\partial L}{\partial\delta} = \begin{pmatrix} \frac{\partial L}{\partial\theta} \\ \frac{\partial L}{\partial\alpha_1} \\ \vdots \\ \frac{\partial L}{\partial\alpha_n} \end{pmatrix} = \sum_{t=1}^{T} \begin{pmatrix} \sum_{i=1}^{n} \frac{\partial\ell_{it}(\theta,\alpha_i)}{\partial\theta} \\ \frac{\partial\ell_{1t}(\theta,\alpha_1)}{\partial\alpha_1} \\ \vdots \\ \frac{\partial\ell_{nt}(\theta,\alpha_n)}{\partial\alpha_n} \end{pmatrix} = \begin{pmatrix} d_\theta \\ d_\alpha \end{pmatrix}$$

$$\frac{\partial^2 L}{\partial\delta\partial\delta'} = \sum_{t=1}^{T} \begin{pmatrix} \sum_{i=1}^{n} \frac{\partial^2\ell_{it}(\theta,\alpha_i)}{\partial\theta\partial\theta'} & \frac{\partial^2\ell_{1t}(\theta,\alpha_1)}{\partial\theta\partial\alpha_1'} & \dots & \frac{\partial^2\ell_{nt}(\theta,\alpha_n)}{\partial\theta\partial\alpha_n'} \\ \frac{\partial^2\ell_{1t}(\theta,\alpha_1)}{\partial\alpha_1\partial\theta'} & \frac{\partial^2\ell_{1t}(\theta,\alpha_1)}{\partial\alpha_1\partial\alpha_1'} & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2\ell_{nt}(\theta,\alpha_n)}{\partial\alpha_n\partial\theta'} & 0 & \dots & \frac{\partial^2\ell_{nt}(\theta,\alpha_n)}{\partial\alpha_n\partial\alpha_n'} \end{pmatrix} = \begin{pmatrix} H_{\theta\theta} & H_{\theta\alpha} \\ H_{\theta\alpha}' & H_{\alpha\alpha} \end{pmatrix}$$

and

$$d_\alpha = \begin{pmatrix} d_{\alpha 1} \\ \vdots \\ d_{\alpha n} \end{pmatrix}, \quad H_{\alpha\alpha} = \begin{pmatrix} H_{\alpha\alpha 1} & & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & H_{\alpha\alpha n} \end{pmatrix}, \quad H_{\theta\alpha} = \begin{pmatrix} H_{\theta\alpha 1} & \dots & H_{\theta\alpha n} \end{pmatrix}$$

so that $d_\theta = \sum_{i=1}^n d_{\theta i}$ and $H_{\theta\theta} = \sum_{i=1}^n H_{\theta\theta i}$, and

$$H_{\theta\alpha} H_{\alpha\alpha}^{-1} = \begin{pmatrix} H_{\theta\alpha 1} H_{\alpha\alpha 1}^{-1} & \dots & H_{\theta\alpha n} H_{\alpha\alpha n}^{-1} \end{pmatrix}$$

$$H_{\theta\alpha} H_{\alpha\alpha}^{-1} H_{\alpha\theta} = \sum_{i=1}^n H_{\theta\alpha i} H_{\alpha\alpha i}^{-1} H_{\alpha\theta i}$$

Letting

$$\begin{pmatrix} H_{\theta\theta} & H_{\theta\alpha} \\ H'_{\theta\alpha} & H_{\alpha\alpha} \end{pmatrix}^{-1} = \begin{pmatrix} H^{\theta\theta} & H^{\theta\alpha} \\ H^{\theta\alpha\prime} & H^{\alpha\alpha} \end{pmatrix}$$

where

$$\begin{aligned} H^{\theta\theta} &= \left(H_{\theta\theta} - H_{\theta\alpha} H_{\alpha\alpha}^{-1} H_{\alpha\theta} \right)^{-1} \\ H^{\theta\alpha} &= -H^{\theta\theta} H_{\theta\alpha} H_{\alpha\alpha}^{-1} \\ H^{\alpha\alpha} &= H_{\alpha\alpha}^{-1} + H_{\alpha\alpha}^{-1} H_{\alpha\theta} H^{\theta\theta} H_{\theta\alpha} H_{\alpha\alpha}^{-1} \end{aligned}$$

the partitioned formula gives:

$$\begin{pmatrix} \Delta\theta \\ \Delta\alpha \end{pmatrix} = - \begin{pmatrix} H^{\theta\theta} & H^{\theta\alpha} \\ H^{\theta\alpha\prime} & H^{\alpha\alpha} \end{pmatrix} \begin{pmatrix} d_\theta \\ d_\alpha \end{pmatrix}$$

or

$$\begin{aligned} \Delta\theta &= -\left(H^{\theta\theta} d_\theta + H^{\theta\alpha} d_\alpha \right) \\ \Delta\alpha &= -\left(H^{\theta\alpha\prime} d_\theta + H^{\alpha\alpha} d_\alpha \right). \end{aligned}$$

We have

$$
\begin{aligned}
H^{\theta\theta} &= \left[\sum_{i=1}^{n}\left(H_{\theta\theta i} - H_{\theta\alpha i}H_{\alpha\alpha i}^{-1}H_{\alpha\theta i}\right)\right]^{-1} \\
H^{\theta\alpha}d_\alpha &= -H^{\theta\theta}H_{\theta\alpha}H_{\alpha\alpha}^{-1}d_\alpha = -H^{\theta\theta}\sum_{i=1}^{n}H_{\theta\alpha i}H_{\alpha\alpha i}^{-1}d_{\alpha i}
\end{aligned}
$$

and

$$
\begin{aligned}
-\Delta\theta &= H^{\theta\theta}d_\theta + H^{\theta\alpha}d_\alpha = H^{\theta\theta}d_\theta - H^{\theta\theta}\sum_{i=1}^{n}H_{\theta\alpha i}H_{\alpha\alpha i}^{-1}d_{\alpha i} \\
&= H^{\theta\theta}\left(d_\theta - \sum_{i=1}^{n}H_{\theta\alpha i}H_{\alpha\alpha i}^{-1}d_{\alpha i}\right)
\end{aligned}
$$

so that

$$
\Delta\theta = -\left[\sum_{i=1}^{n}\left(H_{\theta\theta i} - H_{\theta\alpha i}H_{\alpha\alpha i}^{-1}H_{\alpha\theta i}\right)\right]^{-1}\sum_{i=1}^{n}\left(d_{\theta i} - H_{\theta\alpha i}H_{\alpha\alpha i}^{-1}d_{\alpha i}\right).
$$

Similarly, we have[7]

$$
\Delta\alpha = -H_{\alpha\alpha}^{-1}\left(d_\alpha + H_{\alpha\theta}\Delta\theta\right)
$$

so that

$$
\Delta\alpha_i = -H_{\alpha\alpha i}^{-1}\left(d_{\alpha i} + H_{\alpha\theta i}\Delta\theta\right), \qquad (i = 1, ..., n)
$$

---

[7]Note that

$$
\begin{aligned}
-\Delta\alpha &= H^{\theta\alpha\prime}d_\theta + H^{\alpha\alpha}d_\alpha \\
&= -H_{\alpha\alpha}^{-1}H_{\alpha\theta}H^{\theta\theta}d_\theta + \left(H_{\alpha\alpha}^{-1} + H_{\alpha\alpha}^{-1}H_{\alpha\theta}H^{\theta\theta}H_{\theta\alpha}H_{\alpha\alpha}^{-1}\right)d_\alpha \\
&= H_{\alpha\alpha}^{-1}\left(d_\alpha + H_{\alpha\theta}H^{\theta\theta}H_{\theta\alpha}H_{\alpha\alpha}^{-1}d_\alpha - H_{\alpha\theta}H^{\theta\theta}d_\theta\right) \\
&= H_{\alpha\alpha}^{-1}\left(d_\alpha - H_{\alpha\theta}H^{\theta\alpha}d_\alpha - H_{\alpha\theta}H^{\theta\theta}d_\theta\right) \\
&= H_{\alpha\alpha}^{-1}\left[d_\alpha - H_{\alpha\theta}\left(H^{\theta\alpha}d_\alpha + H^{\theta\theta}d_\theta\right)\right] \\
&= H_{\alpha\alpha}^{-1}\left(d_\alpha + H_{\alpha\theta}\Delta\theta\right).
\end{aligned}
$$

4.B A regression-based iteration

Alternatively, we may consider a Gauss-Newton approach after enforcing block diagonality. The motivation is the same as in Berndt, Hall, Hall, and Hausman (1974) in that the nonzero components of the Hessian are approximated by outer product terms. The advantages of this procedure are that it only requires first derivatives and that it leads to a regression-based iteration.

Let us introduce the notation

$$\begin{aligned} d_{\theta it} &= \frac{\partial \ell_{it}(\theta,\alpha_i)}{\partial \theta}, \quad d_{\alpha it} = \frac{\partial \ell_{it}(\theta,\alpha_i)}{\partial \alpha_i} \\ \Psi_{\theta\theta i} &= \sum_{t=1}^{T} d_{\theta it} d_{\theta it}', \quad \Psi_{\alpha\alpha i} = \sum_{t=1}^{T} d_{\alpha it} d_{\alpha it}', \quad \Psi_{\theta\alpha i} = \sum_{t=1}^{T} d_{\theta it} d_{\alpha it}'. \end{aligned}$$

The Kth step of the iteration of the Gauss–Newton algorithm for obtaining $\widehat{\theta}$ and $\widehat{\alpha}$ takes the form

$$\theta_{[K]} - \theta_{[K-1]} = -\left[\sum_{i=1}^{n}\left(\Psi_{\theta\theta i} - \Psi_{\theta\alpha i}\Psi_{\alpha\alpha i}^{-1}\Psi_{\alpha\theta i}\right)\right]^{-1}\sum_{i=1}^{n}\left(d_{\theta i} - \Psi_{\theta\alpha i}\Psi_{\alpha\alpha i}^{-1}d_{\alpha i}\right)$$

$$\alpha_{i[K]} - \alpha_{i[K-1]} = -\Psi_{\alpha\alpha i}^{-1}\left[d_{\alpha i} + \Psi_{\alpha\theta i}\left(\theta_{[K]} - \theta_{[K-1]}\right)\right], \quad (i = 1, ..., n)$$

where all derivatives are evaluated at $\theta_{[K-1]}$ and $\alpha_{i[K-1]}$.

Thus,

$$\theta_{[K]} - \theta_{[K-1]} = -\left(\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{d}_{\theta it}\widetilde{d}_{\theta it}'\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T}\widetilde{d}_{\theta it}$$

where

$$\widetilde{d}_{\theta it} = d_{\theta it} - \widetilde{\Pi}_i d_{\alpha it}$$

and

$$\widetilde{\Pi}_i = \Psi_{\theta\alpha i}\Psi_{\alpha\alpha i}^{-1},$$

so that the $\widetilde{d}_{\theta it}$ are the residuals of individual-specific regressions of $d_{\theta it}$ on $d_{\alpha it}$. Next, $\theta_{[K]} - \theta_{[K-1]}$ can be calculated as a pooled regression of minus one on $\widetilde{d}_{\theta it}$. Finally, $\alpha_{i[K]} - \alpha_{i[K-1]}$ can be obtained as a regression of $-\left[1 + d'_{\theta it}\left(\theta_{[K]} - \theta_{[K-1]}\right)\right]$ on $d_{\alpha it}$:

$$\alpha_{i[K]} - \alpha_{i[K-1]} = -\left(\sum_{t=1}^{T} d_{\alpha it} d'_{\alpha it}\right)^{-1} \sum_{t=1}^{T} d_{\alpha it}\left[1 + d'_{\theta it}\left(\theta_{[K]} - \theta_{[K-1]}\right)\right], \quad (i = 1, ..., n).$$

General Conclusions

This doctoral thesis considers new models and estimation methods for the analysis of the wage distribution and the labour market histories, from a dynamic perspective. The first chapter studies gender differences in wage growth and job mobility in the initial stages of workers' careers. Similar studies has been conducted on data from US, Italy and Finland, and in chapter 1, I extend the analysis to data from the Spanish section of the European Community Household Panel (1994-2001). First, I propose an experience measure that - as opposed to the conventional potential experience variable - considers the existence of discontinuities in the professional career of workers. Secondly, I analyse gender differences in job mobility patterns among young workers. From the comparison between the proposed experience measure - accumulated experience - and the one used normally - potential experience - it turns out that wage returns to experience are higher with the more accurate measure and that difference is greater for women than men. This result suggests the existence of a gender wage penalty to interruptions. Regarding job changes, the findings indicate that turnover rates are similar for men and women among young workers. Differences come from the side of some characteristics that are relevant for one of the two groups and not for the other, specially in case of promotion or in transitions to non-employment. For men, holding a position with responsibility or having a family it turns out to be important when changing job. On the contrary, for women it is relevant the type of journey or the size of the firm. Finally, in addition to the gender penalty to interruptions, I also find that early-career wage growth is greater for men than for women, and this is specially true in years when job changes occur.

In the second chapter, I consider a model for the heterogeneity and dynamics of the conditional mean and the conditional variance of individual wages. In particular, I propose a dynamic panel data model with individual effects both in the mean and in a conditional ARCH type variance function. I posit a distribution for earning shocks and build a modified likelihood function for estimation and inference in a fixed-T context. Using a newly developed bias-corrected likelihood approach it is possible to reduce the estimation bias to a term of order $1/T^2$. The small sample performance of bias corrected estimators is investigated in a Monte Carlo simulation study. The simulation results show that the bias of the maximum likelihood estimator is substantially corrected for designs that are broadly calibrated to the Panel Study of Income Dynamics (PSID). The empirical analysis is conducted on data drawn from the 1968-1993 PSID. Focusing on US data is interesting because we do not know much how the volatilities of individual wages behave in a period of increasing aggregate inequality. In the data, I find that it is important to account for individual unobserved heterogeneity and dynamics in the variance, and that the latter is driven by job mobility. I also find that the model explains the non-normality observed in logwage data. Finally, the chapter includes an illustration of some implications that ARCH effects would have on consumption growth. The main conclusion is that an increase in individual risk induces a reduction on current consumption, and this effect is more important for the less educated people, slightly significant for the graduate and insignificant for the college educated. This result goes in line with the idea that there are more insurance possibilities for these latter.

The third chapter presents a model that explicitly considers job changes in the dynamics of wages and in the heterogeneity pattern. I propose a specification with two different parameters to capture dynamics within jobs and across jobs, and where the unobserved heterogeneity shows a richer pattern, as well, composed of both individual and job-specific effects. The potential endogeneity of job mobility in relation to earnings is circumvented using an instrument variable estimation method that controls for those unobserved heterogeneity components. In the empirical application, I use data on work histories drawn from the PSID. Regarding results, once we control for individual and job-specific effects, the dynamics within jobs is almost zero, whereas across jobs is significant but small. For the dynamics, the distinction between voluntary and involuntary transitions turns out to be irrelevant. However, that distinction matters in the case of the components of the cross-sectional variance. The estimated variance of the job-specific effects represents around one third of the variance for the individual fixed effects. If I consider a subsample that only includes involuntary job changes, the estimated variance of the heterogeneity across jobs increases up to one half.

A natural next step in my research agenda would be the comparison of the results from chapters 2 and 3 with the corresponding to European countries. Another interesting extension would be the consideration of the labour market participation decision and, thus, the inclusion of women and transitions job-to-nonemployment and nonemployment-to-job into the analysis.

The fourth chapter is mainly a methodological contribution, related to the computational calculation in practice of bias corrections of the type presented in chapter 2. In particular, chapter 4 considers estimation of nonlinear panel data models that include multiple individual fixed effects. Estimation of these models is complicated both by the difficulty of estimating models with possibly thousands of coefficients and also by the incidental parameters problem; that is, noisy estimates of the fixed effects when the time dimension is short contaminate the estimates of the common parameters due to the nonlinearity of the problem. This chapter shows how to use an iterated algorithm which simplifies estimation in a nonlinear model with multiple fixed effects and discusses its application to bias corrected concentrated likelihoods.

This last chapter represents another exciting line of future research, since more results are needed in how well bias corrections methods work for different models and data sets of interest in applied econometrics, and in the theoretical properties that would help us to narrowing the choice between alternative bias reducing estimation methods.

References

Abowd, J.M. and D. Card (1989), "On the Covariance Structure of Earnings and Hours Changes", *Econometrica*, 57, 441-445.

Abowd, J.M. and C. Kang (2002), "Simultaneous Determination of Wage Rates and Tenure", unpublished manuscript.

Abowd, J.M., Kramarz, F. and D.N. Margolis (1999), "High Wage Workers and High Wage Firms", *Econometrica*, 67, 251-333.

Albarrán, P. (2004), "Modelos Econométricos de Varianza Condicional para Datos de Panel", cap. 3, Tesis Doctoral, Universidad Complutense.

Altonji, J. G. and C. H. Paxson (1992), "Labor supply, hours constraints and job mobility", *Journal of Human Resources*, 27, 256-278.

Altonji, J.G. and R. Shakotko (1987), "Do wages rise with job seniority?", *Review of Economic Studies,* 54, 437-459.

Alvarez, J. and M. Arellano (2004), "Robust likelihood estimation of dynamic panel data models", CEMFI Working Paper No. 0421.

Alvarez, J., M. Browning,, and M. Ejrnæs (2001), "Modelling Income Processes with Lots of Heterogeneity", unpublished manuscript.

Amman, R. A. and T. J. Klein (2007), "Returns to Type or Tenure?", IZA Discussion Paper 2773.

Arellano, M. (2003), "Discrete Choices with Panel Data", *Investigaciones Económicas*, 27, 423-458.

Arellano, M. and S. Bond (1991), “Some tests of specification for panel data: Monte Carlo Evidence and an Application to Employment Equations”, *Review of Economic Studies*, 58, 277-297.

Arellano, M. and J. Hahn (2006a), “Understanding bias in nonlinear panel models: Some recent developments”, In: R. Blundell, W. Newey, and T. Persson (eds.): *Advances in Economics and Econometrics, Ninth World Congress*, Cambridge University Press, forthcoming.

Arellano, M. and J. Hahn (2006b), “A likelihood-based approximate solution to the incidental parameter problem in dynamic nonlinear models with multiple effects”, unpublished manuscript.

Arellano, M. and B. Honoré (2001), “Panel Data Models. Some Recent Developments”, in J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*, 5.

Bagger, J. (2007), “Early Career Wage Profiles and Mobility Premiums”, unpublished manuscript.

Banks, J., R. Blundell, and A. Brugiavini (2001), “Risk Pooling, Precautionar Saving and Consumption Growth”, *Review of Economic Studies*, 68, 757-779.

Barron, J. M., D. A. Black and M. A. Loewenstein (1993), “Gender Differences in Training, Capital, and Wages”, *Journal of Human Resources*, 28, 343-364.

Becker, G. S. (1975), “Human Capital: A Theoretical and Empirical Analysis”, Second Edition, New York: NBER and Columbia University Press.

Berndt, E. K., B. H. Hall, R. E. Hall, and J. A. Hausman (1974), “Estimation and Inference in Nonlinear Structural Models”, *Annals of Economic and Social Measurement*, 3-4, 653-665.

Blundell, R., L. Pistaferri, and I. Preston (2005), “Consumption Inequality and Partial Insurance”, The Institute for Fiscal Studies WP04-28.

Bollerslev, T. (1986), “Generalized Autoregressive Conditional Heteroskedasticity”, *Journal of Econometrics*, 31, 307-327.

Bonhomme, S. and G. Jolivet (2006), “The Pervasive Absence of Compensating Differentials”, unpublished manuscript.

Booth, A. L. and M. Francesconi (2000), “Job Mobility in 1990s Britain: Does Gender Matter?”, *Research in Labor Economics*, 19, 173-189.

Booth, A. L., M. Francesconi and J. Frank (2002), “Temporary Jobs: Stepping Stones or Dead Ends?”, *The Economic Journal*, 112, F189-F213.

Brown, J.N. and A. Light (1992), “Interpreting Panel Data on Job Tenure”, Journal of Labor Economics, 10, 219-257.

Brown, R.S., M. Moon, and B.S. Zoloth (1980), “Incorporating Occupational Attainment in Studies of Male-Female Earnings Differentials”, *Journal of Human Resources*, 15, 3-28.

Browning, M. and A. Lusardi (1996), “Household Saving: Micro Theories and Micro Facts”, *Journal of Economic Literature*, XXXIV, 1797-1855.

Burdett, K. (1978), “A Theory of Employer Job Search and Quit Rates”, *American Economic Review*, 68, 212-220.

Caparrós Ruiz, A., Navarro Gomez, M. L., and Rueda Narváez, M. F. (2004), “Gender Wage Gaps and Job Mobility in Spain”, *International Journal of Manpower*, 25, 264-278.

Carro, J. (2006), “Estimating Dynamic Panel Data Discrete Choice Models with Fixed Effects”, *Journal of Econometrics*, in press.

Chamberlain, G. (1980), “Analysis of Covariance with Qualitative Data”, *Review of Economic Studies*, 47, 225-238.

Chamberlain, G. (1984), “Panel Data”, in Z. Griliches and M. D. Intriligator (eds.), *Handbook of Econometrics*, 2.

Chamberlain, G. and K. Hirano (1999), “Predictive Distributions Based on Longitudinal Earnings Data”, *Annales d”Economie et de Statistique*, 55-56, 211-242.

Corcoran, M. E. and G. J. Duncan (1979), “Work History, Labor Force Attachment, and Earnings Differences between the Sexes”, *Journal of Human Resources*, 14, 3-20.

Cox, D. R. and N. Reid (1987), “Parameter Orthogonality and Approximate Conditional Inference” (with discussion), *Journal of the Royal Statistical Society*, Series B, 49, 1-39.

Del Bono, E. and D. Vuri (2006), "Is It the Way She Moves? New Evidence on the Gender Wage Growth Gap in the Early Careers of Men and Women in Italy", IZA Discussion Paper 2523.

De la Rica, S. and A. Ugidos (1995), "¿Son las diferencias en capital humano determinantes de las diferencias salariales observadas entre hombres y mujeres?", *Investigaciones Económicas*, XIX, 395-414.

DiCiccio, T. J. and S. E. Stern (1993), "An adjustment to Profile Likelihood Based on Observed Information", Technical Report, Department of Statistic, Stanford University.

Dolado, J.J. and V. Llorens (2004), "Gender Wage Gaps by Education in Spain: Glass Floors versus Glass Ceilings", CEPR Discussion Paper 4203.

Donohue, J.J. (1988), "Determinants of Job Turnover of Young Men and Women in the United States: A Hazard Rate Analysis". In T. P. Schultz (ed.), *Research in Population Economics*, Vol.6, Greenwich, CT: JAI.

Dustmann, C. and C. Meghir (2005), "Wages, Experience and Seniority", *Review of Economic Studies*, 72, 77-108.

Fernández-Val, I. (2005), "Estimation of Structural Parameters and Marginal Effects in Binary Choice Panel Data Models with Fixed Effects", unpublished manuscript.

García-Crespo, D. (2001), "Promotions in the Spanish labour market: differences by gender", *Oxford Bulletin of Economics and Statistics*, 63, 599-615.

García, J., P. J. Hernández and A. López-Nicolás (2001), "How Wide is the Gap? An Investigation of Gender Wage Differences using Quantile Regression", *Empirical Economics*, 26, 149-168.

Gardeazabal, J. and A. Ugidos (2005), "Gender Wage Discrimination at Quantiles", *Journal of Population Economics*, 18, 165-179.

Gottschalk, P. and R. Moffitt (1994), "The Growth of Earnings Instability in the US Labor Market", *Brookings Paper on Economic Activity*, 2, 217-272.

Greene, W. (2004), "The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects", *The Econometrics Journal*, 7, 98–119.

Gronau, R. (1988), “Sex related wage Differentials and Women”s Interrupted Labor Careers: the Chicken or the Egg”, *Journal of Labor Economics*, 6, 277-301.

Hahn, J. and G. Kuersteiner (2004), “Bias Reduction for Dynamic Nonlinear Panel Models with Fixed Effects”, unpublished manuscript.

Hahn, J. and W. K. Newey (2004), “Jackknife and Analytical Bias Reduction for Nonlinear Panel Models”, *Econometrica*, 72, 1295-1319.

Hall, B. H. (1978), “A General Framework for Time Series - Cross Section Estimation”, *Annales de l’INSEE*, 30-31, 177-202.

Hall, P. and J. L. Horowitz (1996), “Bootstrap Critical Values for Tests Based on Generalized - Method - of - Moments Estimators”, *Econometrica*, 64, 891-916.

Hausman, J. A. and W. E. Taylor (1981), “Panel Data and Unobservable Individual Effects”, *Econometrica*, 49, 1377-1398.

Heckman, J. and B. Singer (1984), “Econometric Duration Analysis”, *Journal of Econometrics*, 27, 63-132.

Hernández, P. J. (1995), “Análisis empírico de la discriminación salarial de la mujer en España”, *Investigaciones Económicas*, XIX, 195-215.

Holtz-Eakin, D., W. Newey, and H. Rosen (1988), “Estimating Vector Autoregressions with Panel Data”, *Econometrica*, 56, 1371-1395.

Horowitz, J. L. (2003), “Bootstrap Methods for Markov Processes”, *Econometrica*, 71, 1049-1082.

Horowitz, J. L. and M. Markatou (1996), “Semiparametric Estimation of Regression Models for Panel Data”, *Review of Economic Studies*, 63, 145-168.

Johnson, W. R. (1978), “A Theory of Job Shopping”, *Quarterly Journal of Economics*, 92, 261-278.

Jovanovic, B. (1979), “Firm-Specific Capital and Turnover”, *Journal of Political Economy*, 87, 1246-1260.

Juhn, C., K. M. Murphy, and B. Pierce (1993), “Wage Inequality and the Rise in Returns to Skill”, *Journal of Political Economy*, 101, 410-442.

Keith, K. and A. McWilliams (1995), “The Wages Effects of Cumulative Job Mobility”, *Industrial and Labor Relations Review*, 49, 121-137.

Lancaster, T. (2002), “Orthogonal Parameters and Panel Data”, *Review of Economic Studies*, 69, 647-666.

Levy, F. and R. J. Murnane (1992), “U.S. Earnings Levels and Earnings Inequality: A Review of Recent Trends and Proposed Explanations”, *Journal of Economic Literature*, 30, 1333-1381.

Light, A. and M. Ureta (1995), “Early-Career Work Experience and Gender Wage Differentials”, *Journal of Labor Economics*, 13, 121-154.

Lillard, L. A. (1999), “Job turnover heterogeneity and person-job-specific time-series wages”, *Annales d’Economie et de Statistique*, 55-56, 183-210.

Lillard, L. A. and R. J. Willis (1978), “Dynamic Aspects of Earnings Mobility”, *Econometrica*, 46, 985-1012.

Lin, C.-C. (2005), “Estimation and Testing of the ARCH(1) effect in Panel Data when N and T are Large”, unpublished manuscript.

Loprest, P.J. (1992), “Gender Differences in Wage Growth and Job Mobility”, *American Economic Review*, 82, 526-532.

Low, H., C. Meghir, and L. Pistaferri (2007), “Wage Risk and Employment Risk over the Life Cycle”, unpublished manuscript.

MaCurdy, T. E. (1982), “The Use of Time-Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis”, *Journal of Econometrics*, 18, 83-114.

Marshall, R. and G. Zarkin (1987), “The effect of job tenure on wage offers”, *Journal of Labor Economics*, 5, 301-324.

Meghir, C. and F. Windmeijer (1999), “Moment conditions for dynamic panel data models with multiplicative individual effects in the conditional variance”, *Annales d’Economie et de Statistique*, 55-56, 317-330.

Meghir, C. and L. Pistaferri (2004), “Income variance dynamics and heterogeneity”, *Econometrica*, 72, 1-32.

Mincer, J. (1974), *Schooling, Experience and Earnings.* Columbia University Press.

Mincer, J. and H. Ofek (1982), "Interrupted Work Careers: Depreciation and Restoration of Human Capital", *Journal of Human Resources*, 17, 3-24.

Mincer, J. and S. Polachek (1974), "Family Investments in Human Capital: Earnings of Women", *Journal of Political Economy*, 82, S76-S108.

Moltó, M. L. (1984), "Estudio empírico de la discriminación de la mujer en el mercado de trabajo académico", *Estadística Española*, 102, 105-128.

Murphy, K. M. and F. Welch (1990), "Empirical Age-Earnings Profiles", *Journal of Labor Economics*, 8, 202-229.

Neal, D. (1995), "Industry-Specific Human Capital: Evidence from Displaced Workers", *Journal of Labor Economics*, 13, 653-677.

Nelson, D. B. (1992), "Conditional Heteroskedasticity in Asset Returns: A New Approach", *Econometrica*, 59, 347-370.

Neumark, D. (1988), "Employers" Discriminatory Behavior and the Estimation of Wage Discrimination", *Journal of Human Resources*, 23, 279-295.

Neyman, J. and E. L. Scott (1948), "Consistent Estimates Based on Partially Consistent Observations", *Econometrica*, 16, 1-32.

Oaxaca, R. (1973), "Male-Female Differentials in Urban Labor Markets", *International Economic Review*, 14, 693-709.

Pace, L. and A. Salvan (2005), "Adjustments of the profile likelihood from a new perspective", *Journal of Statistical Planning and Inference*, forthcoming.

Peinado, A. (1990), "Un análisis microeconómico de las diferencias salariales entre hombres y mujeres", *Información Comercial Española*, 678, 101-109.

Polsky, D. (1999), "Changing Consequences of Job Separation in the United States", *Industrial and Labor Relations Review*, 52, 565-580.

Postel-Vinay, F., J.-M. Robin (2002), "Equilibrium Wage Dispersion with Heterogeneous Workers and Firms", *Econometrica*, 70, 2295-1350.

Rosen, S. (1972), “Learning by experience as joint production”, *Quarterly Journal of Economics*, 86, 366-382.

Royalty, A. B. (1998), “Job-to-Job and Job-to-Nonemployment Turnover by Gender and Education Level”, *Journal of Labor Economics*, 16, 392-443.

Sandell, S. and D. Shapiro (1980), “Work Expectations, Human Capital Accumulation, and the Wages of Women”, *Journal of Human Resources*, 15, 335-353.

Severini , T. A. (1998a), “An Approximation to the Modified Profile Likelihood Function”, *Biometrika*, 85, 403-411.

Severini, T. A. (1998b), “Likelihood Functions for Inference in the Presence of a Nuisance Parameter”, *Biometrika* 85, 507-522.

Severini, T. A. (2002), *Likelihood Methods in Statistics*, Oxford University Press.

Topel, R. (1986), ”Job mobility, search, and earnings growth: A reinterpretation of human capital earnings functions”, *Research in Labor Economics*, 8, 199-233.

Topel, R. (1991), “Specific capital, mobility and wages: Wages rise with job seniority”, *Journal of Political Economy*, 99, 145-176.

Topel, R. and M. Ward (1992), “Job Mobility and the Careers of Young Men”, *Quarterly Journal of Economics*, 107, 441-479.

Wooldridge, J. M. (2001), *Econometric Analysis of Cross Section and Panel Data.* The MIT Press.

Woutersen, T. (2002), “Robustness against Incidental Parameters”, unpublished manuscript.

Appendix A

Tables of Chapter 1

Table A.1: Descriptive Statistics

Observations individual-year	Males Mean	Males SD	Females Mean	Females SD
Number of individuals	543		577	
Number of observations	1537		1726	
Age (last observation)	23.12	3.82	25.09	4.23
Children	0.02	0.13	0.09	0.28
Married	0.06	0.23	0.21	0.41
Hours per week	39.86	7.79	37.29	10.20
Hourly Wage (pta. 1992)	633.45	287.53	601.94	343.58
Accumulated Experience	1.08	1.28	1.04	1.23
Potential Experience	3.86	2.73	4.82	4.18
Tenure	0.75	1.16	0.74	1.08
Years of schooling	12.85	3.45	13.64	3.54
Primary Education	0.26	0.45	0.19	0.39
Graduate Education	0.35	0.47	0.28	0.45
College Education	0.39	0.49	0.53	0.50
Part-time	0.12	0.33	0.25	0.44
Temporary	0.44	0.49	0.40	0.49
Managers and Professionals	0.22	0.42	0.28	0.45
Clerical and Services	0.17	0.38	0.46	0.50
Agriculture and Manufacture	0.37	0.48	0.09	0.29
Unskilled workers	0.24	0.42	0.17	0.37
More than 50 wage-earners	0.31	0.46	0.31	0.46
Between 5 and 49 wage-earners	0.51	0.50	0.41	0.49
Between 1 and 4 wage-earners	0.18	0.39	0.28	0.45

SD: Standard deviation.

Table A.2: Proportion of males/females that work more than a fraction of time

# months	6	18	30	42	54
Males	90.65	84.17	71.22	62.59	38.13
Females	81.17	63.64	46.10	33.70	22.08

Note: fraction of time measured as the number of months employed in the last 5 years.

Table A.3: OLS Regressions by gender. Accumulated Experience

Dependent variable: logwage rate	[I] Males	[I] Females	[II] Males	[II] Females
Accumulated Experience	0.061***	0.059***	0.057***	0.057***
	[0.022]	[0.022]	[0.022]	[0.022]
$(\text{Accumulated Experience})^2$	-0.005	-0.007	-0.005	-0.008*
	[0.004]	[0.005]	[0.004]	[0.005]
Difference Potential -			0.005	-0.008*
Acccumulated Experience			[0.009]	[0.004]
Delayed Entrance to Market			0.034	0.019
			[0.046]	[0.026]
Birth year: 1955-1973	0.177***	0.157***	0.088*	0.182***
	[0.044]	[0.039]	[0.046]	[0.035]
Birth year: 1974-1977	0.080**	0.048	0.065*	0.063**
	[0.033]	[0.035]	[0.034]	[0.028]
Married	-0.041	0.046	-0.033	0.051*
	[0.053]	[0.037]	[0.052]	[0.030]
Children	0.195**	-0.011	0.163	0.005
	[0.097]	[0.046]	[0.100]	[0.043]
Graduate Education	0.077**	0.045*	0.082**	0.026
	[0.034]	[0.027]	[0.033]	[0.029]
College Education	0.170***	0.141***	0.193***	0.151***
	[0.034]	[0.028]	[0.037]	[0.031]
Tenure: < 1 year	-0.039	-0.045	-0.045	-0.044
	[0.037]	[0.032]	[0.037]	[0.032]
Tenure: 1 - 2 years	-0.010	-0.033	-0.016	-0.032
	[0.031]	[0.030]	[0.031]	[0.030]
Civil servant	0.211***	0.208***	0.217***	0.204***
	[0.052]	[0.051]	[0.052]	[0.038]
Temporary	-0.062***	-0.005	-0.059***	-0.007
	[0.025]	[0.027]	[0.025]	[0.024]
Part-time	0.118**	0.103***	0.115**	0.103***
	[0.050]	[0.029]	[0.050]	[0.027]
> 50 wage-earners	0.233***	0.205***	0.230***	0.204***
	[0.041]	[0.031]	[0.041]	[0.026]
5 - 49 wage-earners	0.126***	0.146***	0.124***	0.146***
	[0.036]	[0.027]	[0.036]	[0.024]
Managers and Professionals	0.153***	0.305***	0.161***	0.293***
	[0.045]	[0.044]	[0.045]	[0.039]
Clerical and Services	-0.045	0.030	-0.046	0.023
	[0.034]	[0.033]	[0.033]	[0.030]
Agriculture and Manufacture	0.025	0.034	0.028	0.029
	[0.027]	[0.039]	[0.027]	[0.035]
Constant	5.987***	5.873***	5.967***	5.897***
	[0.070]	[0.087]	[0.070]	[0.080]
Observations	973	1040	973	1040
R^2	0.43	0.47	0.43	0.47

Note: ***, **, * significant at 99%, 95%, 90% level, respectively. Robust standard errors in brackets. Year and region dummies included. Omitted group: Birth year>1977, primary education, more than 2 years of tenure, between 1 and 4 wage-earners, unskilled.

Table A.4: OLS Regressions by gender. Potential Experience

Dependent variable: logwage rate	[I] Males	[I] Females	[II] Males	[II] Females
Potential Experience	0.033***	0.011*	0.033***	0.012*
	[0.011]	[0.006]	[0.011]	[0.006]
(Potential Experience)^2	-0.002*	-0.001***	-0.002*	-0.001***
	[0.001]	[0.000]	[0.001]	[0.000]
Interruptions			0.003	-0.053**
			[0.021]	[0.023]
Birth year: 1955-1973	0.093**	0.187***	0.093**	0.184***
	[0.047]	[0.034]	[0.047]	[0.034]
Birth year: 1974-1977	0.063*	0.068**	0.063*	0.065**
	[0.034]	[0.027]	[0.034]	[0.027]
Married	-0.024	0.057*	-0.023	0.058**
	[0.054]	[0.059]	[0.054]	[0.029]
Children	0.202***	0.010	0.201***	0.013
	[0.101]	[0.043]	[0.101]	[0.043]
Graduate Education	0.085**	0.029	0.085**	0.027
	[0.033]	[0.029]	[0.034]	[0.028]
College Education	0.199***	0.120***	0.200***	0.121***
	[0.037]	[0.031]	[0.037]	[0.031]
Tenure: $<$ 1 year	-0.081**	-0.087***	-0.083**	-0.053
	[0.031]	[0.030]	[0.035]	[0.032]
Tenure: 1 - 2 years	-0.042	-0.058**	-0.044	-0.031
	[0.030]	[0.029]	[0.031]	[0.031]
Civil servant	0.213***	0.197***	0.213***	0.198***
	[0.052]	[0.038]	[0.052]	[0.038]
Temporary	-0.058**	-0.011	-0.058**	-0.008
	[0.025]	[0.024]	[0.025]	[0.024]
Part-time	0.109**	0.094***	0.109**	0.097***
	[0.051]	[0.027]	[0.051]	[0.027]
$>$ 50 wage-earners	0.234***	0.204***	0.234***	0.201***
	[0.032]	[0.026]	[0.041]	[0.026]
5 - 49 wage-earners	0.127***	0.145***	0.127***	0.143***
	[0.030]	[0.024]	[0.037]	[0.024]
Managers and Professionals	0.163***	0.297***	0.163***	0.297***
	[0.045]	[0.039]	[0.045]	[0.039]
Clerical and Services	-0.044	0.024	-0.044	0.023
	[0.034]	[0.030]	[0.034]	[0.030]
Agriculture and Manufacture	0.036	0.040	0.036	0.041
	[0.027]	[0.035]	[0.027]	[0.035]
Constant	5.957***	6.178***	5.957***	5.926***
	[0.075]	[0.083]	[0.075]	[0.080]
Observations	973	1040	973	1.040
R^2	0.43	0.47	0.43	0.47

Note: ***, **, * significant at 99%, 95%, 90% level, respectively. Robust standard errors in brackets. Year and region dummies included. Omitted group: Birth year>1977, primary education, more than 2 years of tenure, between 1 and 4 wage-earners, unskilled.

Table A.5: Effect on the log-wage of a marginal change in the experience

Years of	Accumulated Experience				Potential Experience			
	[I]		[II]		[I]		[II]	
	Males	Females	Males	Females	Males	Females	Males	Females
x=0	0.061***	0.059***	0.057***	0.057***	0.033***	0.011*	0.033***	0.012*
	[0.006]	[0.008]	[0.010]	[0.008]	[0.003]	[0.085]	[0.003]	[0.060]
x=1	0.051***	0.045***	0.047***	0.041***	0.029***	0.009*	0.029***	0.010*
	[0.002]	[0.003]	[0.003]	[0.003]	[0.002]	[0.100]	[0.002]	[0.082]
x=2	0.041***	0.031***	0.037***	0.025***	0.025***	0.007	0.025***	0.008
	[0.002]	[0.007]	[0.002]	[0.008]	[0.001]	[0.162]	[0.001]	[0.119]
x=3	0.031**	0.017	0.027**	0.009	0.021***	0.005	0.021***	0.006
	[0.040]	[0.272]	[0.034]	[0.383]	[0.001]	[0.238]	[0.001]	[0.181]
x=4	0.021	0.003	0.017	-0.007	0.017***	0.003	0.017***	0.004
	[0.366]	[0.943]	[0.311]	[0.833]	[0.001]	[0.363]	[0.001]	[0.286]
Mean	0.049***	0.042***	0.045***	0.037***	0.017***	0.002	0.017***	0.003
	[0.002]	[0.002]	[0.002]	[0.003]	[0.001]	[0.440]	[0.001]	[0.353]

Note: ***, **, * significant at 99%, 95%, 90% level, respectively. p-values in brackets. Mean: average experience value in the sample.

Table A.6: GLS/IV Regressions by gender

Dependent variable: logwage rate	Males	Females
Accumulated Experience	0.068***	0.072***
	[0.026]	[0.026]
Experience squared	-0.006	-0.010***
	[0.003]	[0.004]
Birth year: 1955-1973	0.090	0.216**
	[0.148]	[0.093]
Birth year: 1974-1977	0.046	0.078
	[0.104]	[0.080]
Married	0.037	0.058*
	[0.060]	[0.035]
Children	0.415***	-0.039
	[0.099]	[0.059]
College Education	0.405**	0.292*
	[0.203]	[0.158]
Tenure: $<$ 1 year	-0.028	-0.021
	[0.035]	[0.030]
Tenure: 1 - 2 years	0.020	-0.018
	[0.029]	[0.026]
Civil servant	0.060	0.089**
	[0.052]	[0.045]
Temporary	0.009	-0.006
	[0.023]	[0.021]
Part-time	0.108**	0.149***
	[0.051]	[0.028]
$>$ 50 wage-earners	0.114***	0.095***
	[0.035]	[0.034]
5 - 49 wage-earners	0.074**	0.072***
	[0.032]	[0.028]
Managers and Professionals	-0.029	-0.016
	[0.055]	[0.053]
Clerical and Services	-0.042	-0.039
	[0.042]	[0.045]
Agriculture and Manufacture	0.036	0.014
	[0.031]	[0.048]
Regional unemployment rate	-0.005	-0.003
	[0.004]	[0.004]
Constant	5,967***	5,856***
	[0.127]	[0.136]
Observations	973	1040
Number of individuals	425	459

Note: ***, **, * significant at 99%, 95%, 90% level, respectively. Robust standard errors in brackets. Year dummies included. Omitted group: Birth year>1977, non college education, more than 2 year of tenure, between 1 and 4 wage-earners, unskilled.

Table A.7: OLS Regressions by gender. Mean wage growth

Dependent variable: logwage rate	Males	Females
Potential Experience	0.029***	0.015**
	[0.011]	[0.007]
Potential Experience × Job change	0.015	-0.007
	[0.201]	[0.014]
$(\text{Potential Experience})^2$	-0.001	-0.001***
	[0.001]	[0.000]
$(\text{Potential Experience})^2$ × Job change	-0.001	0.001
	[0.002]	[0.001]
Job change	-0.033	0.004
	[0.052]	[0.048]
Interruptions	0.005	-0.053**
	[0.022]	[0.023]
Birth year: 1955-1973	0.094**	0.184***
	[0.039]	[0.034]
Birth year: 1974-1977	0.064**	0.065**
	[0.029]	[0.027]
Married	-0.024	0.062**
	[0.047]	[0.029]
Children	0.198***	0.002
	[0.077]	[0.043]
Graduate Education	0.087***	0.028
	[0.029]	[0.028]
College Education	0.201***	0.122***
	[0.033]	[0.031]
Tenure: < 1 year	-0.086***	-0.053*
	[0.032]	[0.032]
Tenure: 1 - 2 years	-0.048	-0.031
	[0.031]	[0.031]
Civil servant	0.209***	0.198***
	[0.040]	[0.038]
Temporary	-0.057**	-0.008
	[0.022]	[0.024]
Part-time	0.109**	0.098***
	[0.048]	[0.027]
> 50 wage-earners	0.234***	0.200***
	[0.032]	[0.026]
5 - 49 wage-earners	0.127***	0.143***
	[0.031]	[0.024]
Constant	5,958***	5,926***
	[0.069]	[0.081]
Observations	973	1040
R^2	0.43	0.47

Note: ***, **, * significant at 99%, 95%, 90% level, respectively. Robust standard errors in brackets. Year, region and occupation dummies included. Omitted group: Birth year>1977, primary education, more than 2 years of tenure, between 1 and 4 wage-earners.

Table A.8: Multinomial logit estimates. Males. Without occupations

Dependent variable:		Job to job		Job to nonemployment	
transitions from employed	Promotion	Layoff	Quit	Layoff	Quit
Supervisor	0.449	-0.582	-0.320	0.678	2.076***
	[0.475]	[0.507]	[0.629]	[0.524]	[0.613]
Economic centres	0.321	0.299	-0.373	-0.537*	0.052
	[0.323]	[0.237]	[0.301]	[0.277]	[0.348]
Age	-0.023	0.081*	-0.015	-0.052	-0.133*
	[0.061]	[0.043]	[0.056]	[0.055]	[0.079]
> 50 wage-earners	-0.066	0.011	-0.138	0.817*	-0.394
	[0.463]	[0.352]	[0.450]	[0.456]	[0.487]
5 - 49 wage-earners	-0.404	-0.153	-0.147	0.487	-0.615
	[0.446]	[0.325]	[0.402]	[0.427]	[0.423]
Tenure	-1.183***	-2.442***	-2.672***	-1.308***	-1.083*
	[0.436]	[0.352]	[0.526]	[0.486]	[0.583]
Tenure squared	-0.042	0.380***	0.511***	0.055	0.121
	[0.128]	[0.094]	[0.191]	[0.198]	[0.162]
Family	0.978*	0.580	-0.126	-0.956	-0.218
	[0.528]	[0.501]	[0.823]	[1.087]	[1.199]
Part time	-1.222	-0.506	-0.090	0.686	0.602
	[1.091]	[0.545]	[0.604]	[0.484]	[0.577]
Accumulated Experience	1.535***	1.638***	2.447***	1.257***	-0.340
	[0.401]	[0.392]	[0.588]	[0.478]	[0.640]
Experience squared	-0.145	-0.364***	-0.770***	-0.262*	0.100
	[0.092]	[0.120]	[0.247]	[0.146]	[0.176]
Graduate Education	-0.315	-0.757**	-0.606	-0.544	-0.457
	[0.434]	[0.300]	[0.389]	[0.331]	[0.412]
College Education	0.400	-0.822**	-0.112	-1.016**	-1.538**
	[0.477]	[0.361]	[0.448]	[0.451]	[0.661]
Constant	-1.550	-1.922**	-0.457	0.134	2.184
	[1.275]	[0.904]	[1.124]	[1.106]	[1.487]
Observations	682				

Note: ***, **, * significant at 99%, 95%, 90% level, respectively. Standard errors in brackets. Year dummies included. Omitted group: Primary education, between 1 and 4 wage-earners.

Table A.9: Multinomial logit estimates. Females. Without occupations

Dependent variable:		Job to job		Job to nonemployment	
transitions from employed	Promotion	Layoff	Quit	Layoff	Quit
Supervisor	0.776**	0.588*	0.490	0.090	-0.568
	[0.382]	[0.348]	[0.482]	[0.490]	[1.108]
Economic centres	0.548*	-0.104	0.507	-0.239	-0.109
	[0.285]	[0.227]	[0.313]	[0.274]	[0.469]
Age	-0.091**	-0.012	-0.061	-0.053	-0.073
	[0.046]	[0.034]	[0.049]	[0.042]	[0.075]
$>$ 50 wage-earners	-0.609*	-0.042	0.332	-0.479	-1.344**
	[0.360]	[0.308]	[0.408]	[0.366]	[0.595]
5 - 49 wage-earners	-0.693*	0.163	-0.078	-0.100	-1.555***
	[0.354]	[0.289]	[0.410]	[0.328]	[0.589]
Tenure	-1.807***	-2.845***	-2.220***	-2.288***	-1.772***
	[0.376]	[0.332]	[0.421]	[0.414]	[0.649]
Tenure squared	0.212**	0.484***	0.399***	0.462***	0.222
	[0.096]	[0.085]	[0.112]	[0.112]	[0.177]
Family	0.666*	0.083	0.062	-0.056	0.296
	[0.350]	[0.318]	[0.436]	[0.421]	[0.670]
Part time	-0.430	0.424	0.746*	0.290	0.086
	[0.455]	[0.306]	[0.383]	[0.357]	[0.622]
Accumulated Experience	1.929***	1.899***	1.613***	1.212**	1.253**
	[0.373]	[0.345]	[0.444]	[0.504]	[0.567]
Experience squared	-0.234***	-0.351***	-0.299**	-0.424**	-0.103
	[0.080]	[0.091]	[0.123]	[0.188]	[0.118]
Graduate Education	0.226	-0.076	0.076	-0.346	0.413
	[0.428]	[0.312]	[0.419]	[0.371]	[0.549]
College Education	0.702*	-0.201	0.047	-0.145	-0.878
	[0.405]	[0.302]	[0.421]	[0.360]	[0.714]
Constant	0.060	-0.127	-0.545	1.022	0.236
	[1.058]	[0.808]	[1.144]	[0.960]	[1.671]
Observations	674				

Note: ***, **, * significant at 99%, 95%, 90% level, respectively. Standard errors in brackets. Year dummies included. Omitted group: Primary education, between 1 and 4 wage-earners.

Table A.10: Multinomial logit estimates. Males. With occupations

Dependent variable:		Job to job		Job to nonemployment	
transitions from employed	Promotion	Layoff	Quit	Layoff	Quit
Supervisor	0.288	-0.533	-0.301	0.567	1.887***
	[0.513]	[0.515]	[0.641]	[0.566]	[0.659]
Economic centres	0.428	0.224	-0.430	-0.551*	0.002
	[0.349]	[0.242]	[0.308]	[0.287]	[0.364]
Managers and Professionals	0.576	-0.439	-0.524	-0.344	-0.606
	[0.644]	[0.430]	[0.538]	[0.541]	[0.690]
Clerical and Services	0.407	0.132	0.031	0.415	-0.519
	[0.601]	[0.373]	[0.460]	[0.414]	[0.506]
Agriculture and Manufacture	0.218	-0.178	-0.441	-0.238	-1.357***
	[0.509]	[0.309]	[0.379]	[0.353]	[0.460]
Age	-0.054	0.078*	-0.022	-0.058	-0.143*
	[0.068]	[0.045]	[0.059]	[0.058]	[0.083]
$>$ 50 wage-earners	-0.191	0.210	0.080	0.971**	-0.301
	[0.488]	[0.362]	[0.460]	[0.473]	[0.517]
5 - 49 wage-earners	-0.490	-0.107	-0.124	0.542	-0.659
	[0.458]	[0.329]	[0.405]	[0.431]	[0.436]
Tenure	-1.262***	-2.435***	-2.665***	-1.441***	-1.271**
	[0.455]	[0.357]	[0.532]	[0.532]	[0.598]
Tenure squared	-0.032	0.379***	0.518***	0.044	0.183
	[0.133]	[0.095]	[0.193]	[0.235]	[0.165]
Family	1.035*	0.506	-0.234	-1.000	-0.262
	[0.541]	[0.505]	[0.828]	[1.093]	[1.176]
Part time	-1.206	-0.508	-0.098	0.640	0.541
	[1.107]	[0.545]	[0.606]	[0.490]	[0.584]
Accumulated Experience	1.788***	1.645***	2.500***	1.264***	-0.077
	[0.422]	[0.399]	[0.598]	[0.491]	[0.667]
Experience squared	-0.180*	-0.363***	-0.781***	-0.245*	0.047
	[0.094]	[0.121]	[0.250]	[0.148]	[0.192]
Graduate Education	-0.371	-0.780**	-0.609	-0.710**	-0.306
	[0.473]	[0.306]	[0.395]	[0.351]	[0.428]
College Education	0.432	-0.727*	-0.027	-0.940**	-1.373*
	[0.513]	[0.375]	[0.465]	[0.477]	[0.732]
Constant	-1.346	-1.755*	-0.096	0.365	2.950*
	[1.456]	[0.957]	[1.205]	[1.187]	[1.571]
Observations	663				

Note: ***, **, * significant at 99%, 95%, 90% level, respectively. Standard errors in brackets.
Year dummies included. Omitted group: Primary education, 1 - 4 wage-earners, unskilled.

Table A.11: Multinomial logit estimates. Females. With occupations

Dependent variable:	Job to job			Job to nonemployment	
transitions from employed	Promotion	Layoff	Quit	Layoff	Quit
Supervisor	0.732*	0.738**	0.432	0.192	-0.527
	[0.394]	[0.358]	[0.490]	[0.497]	[1.148]
Economic centres	0.553*	-0.061	0.499	-0.166	0.075
	[0.291]	[0.233]	[0.316]	[0.280]	[0.481]
Managers and Professionals	0.257	-0.872*	0.419	-0.477	-0.457
	[0.610]	[0.447]	[0.624]	[0.554]	[1.588]
Clerical and Services	-0.033	-0.662*	0.505	-0.168	1.370
	[0.536]	[0.374]	[0.536]	[0.457]	[1.111]
Agriculture and Manufacture	-0.135	0.225	-0.282	0.591	1,403
	[0.688]	[0.444]	[0.773]	[0.537]	[1.219]
Age	-0.100**	-0.016	-0.060	-0.050	-0.051
	[0.048]	[0.034]	[0.051]	[0.042]	[0.073]
$>$ 50 wage-earners	-0.602	-0.269	0.565	-0.548	-0.994
	[0.385]	[0.328]	[0.431]	[0.389]	[0.656]
5 - 49 wage-earners	-0.701*	0.061	0.014	-0.129	-1.466**
	[0.363]	[0.297]	[0.419]	[0.338]	[0.614]
Tenure	-1.810***	-2.827***	-2.197***	-2.284***	-1.722**
	[0.377]	[0.334]	[0.421]	[0.415]	[0.672]
Tenure squared	0.216**	0.486***	0.389***	0.461***	0.205
	[0.096]	[0.084]	[0.111]	[0.111]	[0.182]
Family	0.641*	0.108	0.040	-0.078	0.274
	[0.354]	[0.320]	[0.434]	[0.424]	[0.658]
Part time	-0.442	0.385	0.714*	0.265	0.075
	[0.453]	[0.308]	[0.384]	[0.360]	[0.631]
Accumulated Experience	1.901***	1.826***	1.595***	1.170**	1.078*
	[0.377]	[0.347]	[0.444]	[0.505]	[0.598]
Experience squared	-0.230***	-0.339***	-0.290**	-0.417**	-0.057
	[0.081]	[0.091]	[0.121]	[0.187]	[0.128]
Graduate Education	0.137	0.122	-0.057	-0.284	0.322
	[0.446]	[0.331]	[0.442]	[0.386]	[0.566]
College Education	0.540	0.263	-0.208	0.100	-0.719
	[0.473]	[0.363]	[0.496]	[0.417]	[0.759]
Constant	0.329	0.321	-0.864	0.983	-1.496
	[1.201]	[0.891]	[1.284]	[1.058]	[2.000]
Observations	668				

Note: ***, **, * significant at 99%, 95%, 90% level, respectively. Standard errors in brackets.
Year dummies included. Omitted group: Primary education, 1 - 4 wage-earners, unskilled.

Table A.12: Means and standard deviations of the predicted probabilities by gender

		Job change				
		Job to job			To nonemployment	
	To stay	Promotion	Layoff	Quit	Layoff	Quit
Males	0.454	0.088	0.181	0.095	0.121	0.061
	[0.214]	[0.111]	[0.112]	[0.065]	[0.078]	[0.068]
Females	0.418	0.110	0.222	0.088	0.124	0.038
	[0.212]	[0.106]	[0.128]	[0.046]	[0.082]	[0.037]

Note: standard deviations in brackets.

Table A.13: Means of Individual Marginal Effects by gender. To stay

To stay	Males		Females	
	[I]	[II]	[I]	[II]
Supervisor	-0.102	-0.075	-0.091	-0.101
	[0.108]	[0.125]	[0.104]	[0.105]
Economic centres	0.004	0.009	-0.019	-0.028
	[0.095]	[0.101]	[0.083]	[0.085]
Age	0.001	0.003	0.010	0.010
	[0.017]	[0.019]	[0.013]	[0.014]
$>$ 50 wage-earners	-0.012	-0.033	0.058	0.062
	[0.132]	[0.136]	[0.104]	[0.110]
5 - 49 wage-earners	0.029	0.027	0.046	0.050
	[0.126]	[0.131]	[0.101]	[0.104]
Tenure	0.328**	0.333**	0.386***	0.381***
	[0.123]	[0.131]	[0.116]	[0.116]
Family	-0.053	-0.046	-0.039	-0.038
	[0.170]	[0.174]	[0.101]	[0.101]
Part time	0.001	0.010	-0.057	-0.052
	[0.286]	[0.302]	[0.109]	[0.109]
Accumulated Experience	-0.224*	-0.235*	-0.266**	-0.257**
	[0.116]	[0.125]	[0.111]	[0.111]
Graduate Education	0.115	0.120	0.003	-0.005
	[0.145]	[0.158]	[0.119]	[0.123]
College Education	0.117	0.100	-0.001	-0.031
	[0.159]	[0.167]	[0.116]	[0.136]
Managers and Professionals		0.052		0.054
		[0.199]		[0.186]
Clerical and Services		-0.031		0.024
		[0.159]		[0.149]
Agriculture and Manufacture		0.058		-0.055
		[0.146]		[0.169]

Note:***, **, * significant at 99%, 95%, 90%, respectively. Standard Errors in brackets. Omitted group: Primary Education, 1-4 wage-earners, unskilled.

Table A.14: Means of Individual Marginal Effects by gender. Promotion

Promotion	Males		Females	
	[I]	[II]	[I]	[II]
Supervisor	0.021	0.013	0.057	0.045
	[0.034]	[0.034]	[0.045]	[0.044]
Economic centres	0.025	0.031	0.052**	0.049*
	[0.022]	[0.022]	[0.025]	[0.026]
Age	-0.002	-0.004	-0.007*	-0.007*
	[0.004]	[0.004]	[0.004]	[0.004]
> 50 wage-earners	-0.009	-0.025	-0.053	-0.050
	[0.035]	[0.035]	[0.034]	[0.036]
5 - 49 wage-earners	-0.024	-0.030	-0.061*	-0.060*
	[0.029]	[0.029]	[0.031]	[0.032]
Tenure	-0.022	-0.024	-0.039	-0.039
	[0.020]	[0.019]	[0.023]	[0.021]
Family	0.085	0.089	0.063	0.063
	[0.063]	[0.063]	[0.040]	[0.040]
Part time	-0.051	-0.051	-0.055*	-0.054*
	[0.035]	[0.035]	[0.029]	[0.029]
Accumulated Experience	0.056***	0.066***	0.088***	0.088***
	[0.021]	[0.022]	[0.024]	[0.024]
Graduate Education	0.000	-0.003	0.025	0.011
	[0.029]	[0.029]	[0.041]	[0.040]
College Education	0.066*	0.055	0.077**	0.047
	[0.037]	[0.037]	[0.039]	[0.042]
Managers and Professionals		0.061		0.049
		[0.056]		[0.061]
Clerical and Services		0.024		0.005
		[0.044]		[0.046]
Agriculture and Manufacture		0.027		-0.029
		[0.035]		[0.052]

Note:***, **, * significant at 99%, 95%, 90%, respectively. Standard Errors in brackets. Omitted group: Primary Education, 1-4 wage-earners, unskilled.

Table A.15: Means of Individual Marginal Effects by gender. Layoff (Job to job)

Layoff	Males		Females	
(Job to job)	[I]	[II]	[I]	[II]
Supervisor	-0.112***	-0.103**	0.061	0.088
	[0.039]	[0.044]	[0.077]	[0.087]
Economic centres	0.057	0.049	-0.035	-0.030
	[0.037]	[0.039]	[0.042]	[0.043]
Age	0.015**	0.016**	0.004	0.003
	[0.007]	[0.008]	[0.006]	[0.007]
> 50 wage-earners	-0.005	0.017	0.029	-0.011
	[0.052]	[0.056]	[0.050]	[0.053]
5 - 49 wage-earners	-0.012	-0.006	0.064	0.045
	[0.050]	[0.050]	[0.054]	[0.056]
Tenure	-0.181***	-0.175***	-0.223***	-0.217***
	[0.042]	[0.044]	[0.047]	[0.050]
Family	0.089	0.082	-0.010	-0.003
	[0.105]	[0.106]	[0.053]	[0.055]
Part time	-0.072	-0.073	0.049	0.045
	[0.069]	[0.072]	[0.062]	[0.062]
Accumulated Experience	0.097**	0.092**	0.138***	0.130***
	[0.038]	[0.041]	[0.040]	[0.041]
Graduate Education	-0.066	-0.068	-0.013	0.023
	[0.041]	[0.044]	[0.053]	[0.061]
College Education	-0.083	-0.076	-0.043	0.033
	[0.051]	[0.055]	[0.056]	[0.067]
Managers and Professionals		-0.046		-0.127*
		[0.063]		[0.074]
Clerical and Services		0.005		-0.118
		[0.060]		[0.079]
Agriculture and Manufacture		0.001		0.007
		[0.050]		[0.095]

Note:***, **, * significant at 99%, 95%, 90%, respectively. Standard Errors in brackets. Omitted group: Primary Education, 1-4 wage-earners, unskilled.

Table A.16: Means of Individual Marginal Effects by gender. Quit (Job to job)

Quit	Males		Females	
(Job to job)	[I]	[II]	[I]	[II]
Supervisor	-0.045	-0.041	0.015	0.006
	[0.030]	[0.034]	[0.042]	[0.040]
Economic centres	-0.033	-0.037	0.039	0.036
	[0.025]	[0.026]	[0.025]	[0.024]
Age	-0.001	-0.002	-0.003	-0.003
	[0.005]	[0.005]	[0.004]	[0.004]
> 50 wage-earners	-0.017	-0.006	0.047	0.051
	[0.033]	[0.036]	[0.035]	[0.039]
5 - 49 wage-earners	-0.007	-0.006	0.004	0.012
	[0.035]	[0.036]	[0.027]	[0.028]
Tenure	-0.100***	-0.096***	-0.049**	-0.051**
	[0.028]	[0.029]	[0.021]	[0.021]
Family	-0.023	-0.029	-0.005	-0.007
	[0.052]	[0.051]	[0.032]	[0.031]
Part time	-0.009	-0.009	0.057	0.056
	[0.053]	[0.055]	[0.041]	[0.041]
Accumulated Experience	0.087***	0.089***	0.038*	0.041*
	[0.029]	[0.031]	[0.023]	[0.023]
Graduate Education	-0.021	-0.019	0.008	-0.007
	[0.029]	[0.031]	[0.033]	[0.033]
College Education	0.026	0.033	0.004	-0.028
	[0.040]	[0.044]	[0.033]	[0.039]
Managers and Professionals		-0.029		0.060
		[0.040]		[0.066]
Clerical and Services		-0.007		0.054
		[0.037]		[0.044]
Agriculture and Manufacture		-0.022		-0.034
		[0.030]		[0.043]

Note:***, **, * significant at 99%, 95%, 90%, respectively. Standard Errors in brackets. Omitted group: Primary Education, 1-4 wage-earners, unskilled.

Table A.17: Means of Individual Marginal Effects by gender. Layoff (To nonemployment)

Layoff	Males		Females	
(Job to nonemployment)	[I]	[II]	[I]	[II]
Supervisor	0.037	0.033	-0.023	-0.018
	[0.057]	[0.060]	[0.046]	[0.049]
Economic centres	-0.058**	-0.056*	-0.032	-0.026
	[0.028]	[0.029]	[0.030]	[0.031]
Age	-0.005	-0.005	-0.003	-0.003
	[0.006]	[0.006]	[0.005]	[0.005]
> 50 wage-earners	0.076*	0.079*	-0.036	-0.041
	[0.044]	[0.044]	[0.034]	[0.037]
5 - 49 wage-earners	0.055	0.059	0.000	-0.001
	[0.037]	[0.037]	[0.039]	[0.041]
Tenure	-0.026	-0.035	-0.064**	-0.064**
	[0.030]	[0.031]	[0.031]	[0.032]
Family	-0.080*	-0.079*	-0.018	-0.020
	[0.045]	[0.044]	[0.040]	[0.040]
Part time	0.099	0.092	0.009	0.008
	[0.103]	[0.097]	[0.040]	[0.040]
Accumulated Experience	0.037	0.031	-0.006	-0.008
	[0.032]	[0.032]	[0.034]	[0.035]
Graduate Education	-0.022	-0.035	-0.036	-0.033
	[0.033]	[0.032]	[0.036]	[0.037]
College Education	-0.066*	-0.061	-0.013	0.001
	[0.036]	[0.039]	[0.039]	[0.045]
Managers and Professionals		-0.016		-0.028
		[0.051]		[0.056]
Clerical and Services		0.040		-0.005
		[0.049]		[0.052]
Agriculture and Manufacture		-0.003		0.051
		[0.036]		[0.083]

Note:***, **, * significant at 99%, 95%, 90%, respectively. Standard Errors in brackets. Omitted group: Primary Education, 1-4 wage-earners, unskilled.

Table A.18: Means of Individual Marginal Effects by gender. Quit (To nonemployment)

Quit	Males		Females	
(Job to nonemployment)	[I]	[II]	[I]	[II]
Supervisor	0.200**	0.172**	-0.020	-0.020
	[0.081]	[0.081]	[0.017]	[0.018]
Economic centres	0.006	0.004	-0.005	0.000
	[0.020]	[0.021]	[0.014]	[0.014]
Age	-0.008	-0.008	-0.001	-0.001
	[0.005]	[0.005]	[0.002]	[0.002]
> 50 wage-earners	-0.033	-0.031	-0.046***	-0.031
	[0.029]	[0.030]	[0.018]	[0.020]
5 - 49 wage-earners	-0.040	-0.044	-0.053***	-0.046***
	[0.028]	[0.029]	[0.017]	[0.017]
Tenure	0.001	-0.003	-0.010	-0.009
	[0.024]	[0.025]	[0.012]	[0.013]
Family	-0.017	-0.017	0.005	0.005
	[0.052]	[0.051]	[0.022]	[0.021]
Part time	0.041	0.036	-0.003	-0.003
	[0.057]	[0.054]	[0.018]	[0.018]
Accumulated Experience	-0.053	-0.043	0.008	0.006
	[0.028]	[0.028]	[0.012]	[0.013]
Graduate Education	-0.007	0.004	0.014	0.010
	[0.023]	[0.025]	[0.017]	[0.017]
College Education	-0.055**	-0.050*	-0.024	-0.022
	[0.023]	[0.027]	[0.016]	[0.017]
Managers and Professionals		-0.022		-0.009
		[0.032]		[0.037]
Clerical and Services		-0.032		0.041
		[0.022]		[0.031]
Agriculture and Manufacture		-0.060***		0.060
		[0.022]		[0.081]

Note:***, **, * significant at 99%, 95%, 90% level, respectively. Standard errors in brackets. Omitted group: Primary Education, 1-4 wage-earners, unskilled.

Appendix B

Tables of Chapter 2

Table B.1: AR(1) with fixed effects. Properties of $\hat{\alpha}$ ($T = 8$)

	$\alpha = 0.5$			$\alpha = 0.8$		
Estimator of α	Mean	SD	Mean SE	Mean	SD	Mean SE
MLE	0.2947	0.0173	0.0160	0.5263	0.0163	0.0156
Expected Quantities	0.4365	0.0149	0.0151	0.6541	0.0146	0.0143
Bootstrap Variance	0.4745	0.0213	0.0193	0.7158	0.0182	0.0170
Trimming	0.3782	0.0177	0.0197	0.5986	0.0158	0.0165
Lancaster	0.5006	0.0205	0.0197	0.7989	0.0240	0.0240

Note: N=500; simulations=300; parametric bootstrap samples=300; non parametric bootstrap samples=100; trimming=2. SD: Sample standard deviation. Mean SE: Mean of estimated standard errors by individual block-bootstrap.

Table B.2: AR(1) with fixed effects. Properties of $\hat{\alpha}$ ($T = 16$)

	$\alpha = 0.5$			$\alpha = 0.8$		
Estimator of α	Mean	SD	Mean SE	Mean	SD	Mean SE
MLE	0.4008	0.0109	0.0106	0.6653	0.0097	0.0093
Expected Quantities	0.4589	0.0109	0.0111	0.7093	0.0096	0.0093
Bootstrap Variance	0.4962	0.0119	0.0115	0.7781	0.0106	0.0104
Trimming	0.4577	0.0106	0.0101	0.7093	0.0092	0.0089
Lancaster	0.4999	0.0119	0.0117	0.7993	0.0124	0.0119

Note: N=500; simulations=300; parametric bootstrap samples=200; non parametric bootstrap samples=100; trimming=2. SD: Sample standard deviation. Mean SE: Mean of estimated standard errors by individual block-bootstrap.

Table B.3: AR(1) with multiple fixed effects. Properties of $\hat{\alpha}$ for $\alpha = 0.5$

	$T = 8$		$T = 16$	
Estimator of α	Mean	SD	Mean	SD
MLE	0.2575	0.0169	0.3904	0.0113
Expected Quantities (1st)	0.4214	0.0225	0.4862	0.0160
Expected Quantities (2nd)	0.5115	0.0243	0.5119	0.0157
Bootstrap Variance (1st)	0.3753	0.0442	0.4707	0.0167
Bootstrap Variance (2nd)	0.4336	0.0515	0.4925	0.0172
Trimming	0.3105	0.0467	0.4444	0.0121

Note: N=500; simulations=300; parametric bootstrap samples=300; trimming=2. SD: Sample standard deviation.

Table B.4: AR(1)-EARCH(1) with fixed effects. Properties of $\hat{\alpha}, \hat{\beta}$ for $\alpha = 0.5, \beta = 0.5$

Estimator of $(\alpha, \beta)'$	$T = 8$				$T = 16$			
	Mean $\hat{\alpha}$	SD $\hat{\alpha}$	Mean $\hat{\beta}$	SD $\hat{\beta}$	Mean $\hat{\alpha}$	SD $\hat{\alpha}$	Mean $\hat{\beta}$	SD $\hat{\beta}$
MLE	0.4994	0.0126	−0.1022	0.0845	0.5009	0.0069	0.3670	0.0284
Trimming	0.5012	0.0136	−0.0252	0.0973	0.5009	0.0070	0.4596	0.0284

Note: N=1000; simulations=100; trimming=2. SD: Sample standard deviation. T=8: trimming: 95% successful convergence. T=16: trimming: 100% successful convergence.

Table B.5: AR(1)-EARCH(1) with multiple fixed effects. Properties of $\hat{\alpha}, \hat{\beta}$ for $\alpha = 0.5\,(T = 16)$

Estimator of $(\alpha, \beta)'$	Mean $\hat{\alpha}$	SD $\hat{\alpha}$	Mean $\hat{\beta}$	SD $\hat{\beta}$
$\beta = 0.5$				
MLE	0.3958	0.0092	0.4308	0.0388
Trimming	0.4308	0.0388	0.4819	0.0643
$\beta = 0.0$				
MLE	0.3823	0.0175	−0.0465	0.0077
Trimming	0.4426	0.0210	−0.0286	0.0477

Note: N=1000; simulations=20; trimming=2; trimming $\beta = 0.5$: 85% successful convergence; trimming $\beta = 0.0$: 70% successful convergence. SD: Sample standard deviation.

Table B.6: My sample *vs.* Meghir and Pistaferri (2004)

Number of individuals	Meghir & Pistaferri (2004)		Hospido (2007)		Difference
Starting point		53,013		53,005	8
Latino subsample	(10,022)	42,991	(10,022)	42,983	8
Never Heads	(26,962)	16,029	(26,945)	16,038	−9
Heads, Age, N>9	(11,490)	4,539	(10,791)	5,247	−708
Female	(876)	3,663	(1,211)	4,036	−373
Self-employment, missing	(1323)	2,340	(1,831)	2,205	135
Missing education and race	(187)	2,153	(57)	2,148	5
Outlying wages	(84)	2,069	(82)	2,066	3
Final Sample: Individuals		2,069		2,066	
Final Sample: Observations		31,631		32,066	

Table B.7: Sample 1. Distribution of observations by year

Year	Number of observations	Year	Number of observations
1968	655	1981	1419
1969	694	1982	1464
1970	738	1983	1506
1971	780	1984	1559
1972	856	1985	1626
1973	943	1986	1583
1974	1018	1987	1536
1975	1098	1988	1486
1976	1178	1989	1434
1977	1229	1990	1392
1978	1263	1991	1348
1979	1310	1992	1315
1980	1380	1993	1256

Table B.8: Sample 1. Distribution of observations by education

	Number of Individuals			
Number of Years	Whole sample	High School Dropout	High School Graduate	College Graduate
9	212	52	128	32
10	200	43	122	35
11	155	43	82	30
12	143	36	81	26
13	143	34	87	22
14	147	35	86	26
15	145	38	82	25
16	118	26	71	21
17	127	30	76	21
18	87	20	48	19
19	97	21	57	19
20	91	19	54	18
21	91	25	48	18
22	78	19	44	15
23	52	12	33	7
24	46	15	19	12
25	42	12	27	3
26	52	26	46	20

Table B.9: Sample 1. Descriptive Statistics

	1968	1980	1993
Age	36.99	36.61	41.45
	(6.58)	(9.22)	(5.74)
HS Dropout	0.44	0.25	0.12
HS Graduate	0.41	0.55	0.60
Hours	2272	2153	2135
	(573)	(525)	(560)
Married	0.84	0.83	0.83
White	0.68	0.66	0.69
Children	2.80	1.39	1.36
	(2.06)	(1.28)	(1.23)
Family Size	4.90	3.53	3.51
	(2.01)	(1.58)	(1.45)
North-East	0.18	0.16	0.16
North-Central	0.27	0.25	0.23
South	0.39	0.42	0.44
SMSA	0.68	0.67	0.53

Note: Standard deviations of non-binary variables in parentheses.

Table B.10: Sample 1. α and β estimates

Estimator of $(\alpha, \beta)'$	$\widehat{\alpha}$	$\widehat{\beta}$
MLE	0.4822	0.4832
	(0.0114)	(0.0541)
Trimming ($r = 2$)	0.5690	0.5790
	(0.0397)	(0.0915)

Note: Mean of estimated standard errors by individual block-bootstrap in brackets.

Table B.11: Correlations with observed variables

Dependent variable: $\hat{\psi}_i$	[1]	[2]	[3]	[4]	[5]
Constant	-0.6933	-0.7242	-1.6997	-1.2663	-1.0868
	(0.1784)	(0.1834)	(0.1935)	(0.2142)	(0.2211)
Married	-0.4657	-0.4168	-0.4415	-0.3634	-0.3476
	(0.0683)	(0.0673)	(0.0649)	(0.0640)	(0.0632)
Age	-0.0138	-0.0146	-0.0054	-0.0052	-0.0042
	(0.0038)	(0.0037)	(0.0037)	(0.0040)	(0.0039)
White	-0.5984	-0.4237	-0.4409	-0.5229	-0.4337
	(0.0632)	(0.0651)	(0.0631)	(0.0609)	(0.0617)
Technical Workers		-0.4394	-0.4905		-0.4467
		(0.0912)	(0.0881)		(0.0974)
Administrators		-0.4222	-0.4751		-0.4743
		(0.0943)	(0.0911)		(0.0932)
Sales workers		0.2137	0.2325		0.1712
		(0.1076)	(0.1038)		(0.1015)
Services workers		0.3212	0.2837		0.1761
		(0.0983)	(0.0949)		(0.0921)
Operatives workers		0.0812	0.0919		0.0357
		(0.0886)	(0.0854)		(0.0828)
Movers			0.8177	0.5579	0.5734
			(0.0658)	(0.0686)	(0.0678)
Dropout				0.5319	0.1707
				(0.0872)	(0.1019)
Graduate				0.2260	-0.0511
				(0.0687)	(0.0790)
Tenure: 1-2 years				0.0132	-0.0009
				(0.1493)	(0.1474)
Tenure: 2-3 years				-0.1695	-0.1364
				(0.1175)	(0.1163)
Tenure: 4-9 years				-0.4308	-0.3991
				(0.0978)	(0.0969)
Tenure: 9-19 years				-0.8309	-0.7984
				(0.0929)	(0.0919)
Tenure: 20 years or more				-0.9530	-0.9001
				(0.1008)	(0.1002)
R^2	0.0822	0.1192	0.1808	0.2155	0.2389

Note: Number of observations=2066 individuals. Standard errors in brackets. Omitted group: Craftsman workers, Stayers, Education Colllege, Tenure less than a year.

Table B.12: Meghir and Windmeijer (1999). α and β estimates

Estimator of $(\alpha, \beta)'$	$\widehat{\alpha}$	$\widehat{\beta}$
MLE	0.4904	0.3713
	(0.0099)	(0.0313)
Trimming ($r = 2$)	0.5432	0.4145
	(0.0095)	(0.0337)

Note: Mean of estimated standard errors by individual block-bootstrap in brackets.

Table B.13: Main Descriptive Statistics of the following Distributions

Individual means	Data	Model	Counterfactual 1
Mean	-0.0180	-0.0059	-0.0086
Standard deviation	0.7848	0.8404	0.3876
Individual logvariances	**Data**	**Model**	**Counterfactual 2**
Mean	-1.5980	-1.7054	-1.6703
Standard deviation	1.2762	1.4346	0.7890

Table B.14: Mean elasticities with respect to y_{it-1} at different quantiles

τ	All individuals	$\hat{\psi}_i$ under the mean	$\hat{\psi}_i$ above the mean
0.10	0.5595	0.5752	0.5421
0.20	0.5628	0.5731	0.5513
0.30	0.5651	0.5715	0.5580
0.40	0.5671	0.5702	0.5637
0.50	0.5690	0.5690	0.5690
0.60	0.5709	0.5678	0.5743
0.70	0.5729	0.5665	0.5800
0.80	0.5752	0.5649	0.5867
0.90	0.5785	0.5628	0.5959

Table B.15: Mean marginal effects with respect to past shocks at different quantiles

τ	All individuals	$\hat{\psi}_i$ under the mean	$\hat{\psi}_i$ above the mean
With respect to ϵ_{it-1}			
0.10	0.2543	0.1349	0.3866
0.20	0.2589	0.1346	0.3966
0.30	0.2622	0.1344	0.4039
0.40	0.2650	0.1342	0.4101
0.50	0.2677	0.1340	0.4158
0.60	0.2703	0.1338	0.4216
0.70	0.2732	0.1336	0.4278
0.80	0.2765	0.1334	0.4351
0.90	0.2811	0.1331	0.4451
With respect to ϵ_{it-2}			
0.10	0.1455	0.0726	0.2262
0.20	0.1452	0.0725	0.2258
0.30	0.1451	0.0725	0.2255
0.40	0.1449	0.0724	0.2253
0.50	0.1448	0.0724	0.2251
0.60	0.1447	0.0723	0.2248
0.70	0.1445	0.0723	0.2246
0.80	0.1444	0.0722	0.2243
0.90	0.1441	0.0721	0.2239

Table B.16: Sample 2. Distribution of observations by year

Year	Number of observations	Year	Number of observations
1968	366	1981	708
1969	414	1982	767
1970	446	1983	809
1971	475	1984	858
1972	509	1985	921
1973	543	1986	894
1974	580	1987	866
1975	613	1988	837
1976	645	1989	808
1977	630	1990	766
1978	627	1991	734
1979	644	1992	696
1980	676	1993	653

Table B.17: Sample 2. Distribution of observations by education

	Number of Individuals			
Number of Years	Whole sample	High School Dropout	High School Graduate	College Graduate
9	264	78	133	53
10	182	42	103	37
11	150	31	87	32
12	150	33	88	29
13	131	44	69	18
14	97	29	56	12
15	85	27	43	15
16	64	18	34	12
17	54	13	31	10
18	25	6	13	6
19	38	9	19	10
20	21	4	14	3
21	18	7	8	3
22	20	4	15	1
23	14	4	7	3
24	6	2	3	1
25	17	5	10	2
26	10	3	5	2

Table B.18: Sample 2. Descriptive Statistics

	1968	1980	1993
Age	38.18	39.34	42.60
	(6.35)	(9.24)	(5.65)
HS Dropout	0.43	0.31	0.13
HS Graduate	0.41	0.51	0.62
Hours	2252	2146	2130
	(514)	(483)	(521)
Married	0.83	0.84	0.86
White	0.69	0.66	0.67
Children	2.88	1.39	1.37
	(2.06)	(1.28)	(1.28)
Family Size	5.03	3.65	3.60
	(2.00)	(1.64)	(1.47)
North-East	0.17	0.16	0.16
North-Central	0.29	0.27	0.24
South	0.38	0.45	0.45
SMSA	0.68	0.64	0.52

Note: Standard deviations of non-binary variables in parentheses.

Table B.19: Sample 2. α and β estimates

Estimator of $(\alpha, \beta)'$	$\widehat{\alpha}$	$\widehat{\beta}$
MLE	0.3768	0.0642
	(0.0158)	(0.0846)
Trimming ($r = 2$)	0.4569	0.0757
	(0.0361)	(0.0592)

Note: Mean of estimated standard errors by individual block-bootstrap in brackets.

Table B.20: Attrition. α and β estimates

Estimator of $(\alpha, \beta)'$	$\widehat{\alpha}$	$\widehat{\beta}$
MLE	0.5659	0.5245
	(0.0114)	(0.0412)
Trimming ($r = 2$)	0.6056	0.5693
	(0.0347)	(0.0717)

Note: Mean of estimated standard errors by individual block-bootstrap in brackets.

Table B.21: Consumption Growth Equation

	[1]	[2]	[3]	[4] Total effect
Age	-0.0004	-0.0003	-0.0003	
	(0.0008)	(0.0008)	(0.0008)	
ΔChildren	0.1513	0.1527	0.1527	
	(0.0100)	(0.0101)	(.0101049)	
ΔAdults	0.1593	0.1594	0.1592	
	(0.0119)	(0.0119)	(0.0119)	Dropout
$\pi_{it}^2 \sigma_{it+1}^2$		0.0801	0.1267	0.127 [0.009]
		(0.0308)	(0.0485)	Graduate
$\pi_{it}^2 \sigma_{it+1}^2 \times$			-0.0714	0.055 [0.082]
Graduate			(0.0597)	College
$\pi_{it}^2 \sigma_{it+1}^2 \times$			-0.1089	0.018 [0.905]
College			(0.1562)	
# Obs.		13,723		

Note: clustered standard errors in brackets. Time and cohort dummies included. t-ratios in squared brackets.

Appendix C

Tables of Chapter 3

Table C.1: Distribution of observations by year

Year	Number of observations	Year	Number of observations
1968	613	1981	1,287
1969	668	1982	1,330
1970	726	1983	1,343
1971	762	1984	1,393
1972	815	1985	1,451
1973	885	1986	1,400
1974	965	1987	1,353
1975	1,046	1988	1,302
1976	1,072	1989	1,258
1977	1,104	1990	1,205
1978	1,146	1991	1,173
1979	1,201	1992	1,096
1980	1,251		

Table C.2: Distribution of individuals by number of observations

Number of Years	Number of Individuals	Number of Years	Number Individuals
8	245	17	84
9	211	18	84
10	153	19	79
11	179	20	68
12	143	21	54
13	151	22	35
14	150	23	41
15	130	24	32
16	112	25	62

Table C.3: Descriptive Statistics

	1968	1980	1992
Age	37.16	36.58	40.48
	(6.33)	(8.82)	(5.70)
HS Dropout	0.45	0.26	0.12
HS Graduate	0.40	0.55	0.61
Hours	2,272	2,149	2,197
	(524)	(502)	(489)
Married	0.74	0.80	0.86
White	0.66	0.64	0.69
# Children	2.83	1.45	1.44
	(2.08)	(1.32)	(1.19)
Family Size	4.95	3.60	3.56
	(2.03)	(1.66)	(1.38)
North-East	0.18	0.16	0.17
North-Central	0.26	0.24	0.23
South	0.42	0.46	0.44
SMSA	0.69	0.66	0.54

Note: Standard deviations of non-binary variables in parentheses.

Table C.4: Distribution of Individuals over Jobs by Birth Cohort (percent)

	Maximum Number of jobs							
	1	2	3	4	5	6	<6	N
All	37.70	22.16	17.09	9.34	6.01	3.73	3.97	2,013
Before 1941	51.12	21.67	13.60	5.53	3.29	2.24	2.55	669
1941 on	31.03	22.40	18.82	11.24	7.37	4.46	4.68	1,344

Note: Percentages are computed on total number of individuals in the sample, N. Each cell represents the proportion of individuals who had at most x number of jobs.

Table C.5: Sample Correlations across Individuals

Correlations	No-change at time t	Job loss at time t	Job quit at time t
(w_{it-3}, w_{it-2})	0.902	0.863	0.873
(w_{it-2}, w_{it-1})	0.905	0.689	0.853
(w_{it-1}, w_{it})	0.903	0.590	0.748
(w_{it}, w_{it+1})	0.886	0.816	0.893
(w_{it+1}, w_{it+2})	0.866	0.817	0.857

Table C.6: Sample Wage Annual Growth

Wage growth	Within job	Job loss	Job quit
All	0.010	-0.101	0.035
	(0.271)	(0.716)	(0.444)
Wokers<35	0.021	-0.022	0.075
years old	(0.268)	(0.725)	(0.398)
Wokers≥35	0.002	-0.164	0.013
years old	(0.272)	(0.703)	(0.503)
Dropout	0.001	-0.113	0.012
	(0.332)	(0.810)	(0.506)
Graduate	0.009	-0.098	0.076
	(0.258)	(0.661)	(0.480)
College	0.024	-0.074	0.1022
	(0.203)	(0.680)	(0.457)

Note: standard deviation in parentheses.

Table C.7: Logwages on number of jobs

Number	OLS			Fixed effects		
of jobs	All	Voluntary movers	Involuntary movers	All	Voluntary movers	Involuntary movers
2	-0.017	0.069	-0.371	0.020	0.082	-0.177
	(−1.89)	(7.13)	(−20.45)	(2.84)	(10.80)	(−11.90)
3	-0.044	0.128	-0.442	0.068	0.195	-0.146
	(−3.94)	(9.00)	(−16.23)	(7.45)	(17.92)	(−6.93)
4	-0.076	0.146	-0.597	0.074	0.227	-0.329
	(−5.18)	(7.13)	(−9.86)	(6.29)	(14.92)	(−7.81)
5	-0.139	0.119	-0.740	0.076	0.258	-0.338
	(−6.67)	(3.45)	(−6.25)	(5.07)	(11.66)	(−4.68)
6	-0.175	0.282	-1.305	0.118	0.356	-1.166
	(−6.41)	(5.42)	(−2.41)	(6.31)	(11.00)	(−4.37)
7	-0.391	0.110	-0.997	0.028	0.268	-0.841
	(−9.62)	(1.44)	(−1.84)	(1.07)	(6.03)	(−3.15)

Note: t-ratios in parentheses. All regressions include age and time dummies.

Table C.8: Autorregresive Model of Earnings

	OLS levels	OLS dif	WG	GMM1	GMM2	GMM System	GMM2 AR(2)
y_{it-1}	0.792	-0.313	0.389	0.331	0.321	0 .431	0.329
	(0.009)	(0.012)	(0.006)	(0.023)	(0.022)	(0.020)	(0.025)
y_{it-2}							0.048
							(0.014)
m1	-	-	-	17.25*	-14.55*	-15.58*	-13.83*
m2	-	-	-	2.09*	1.80	2.91*	-0.15
Sargan test	-	-	-	-	304.02	369.18*	295.48
(df)					(275)	(298)	(273)

Note: Heteroskedasticity robust standard errors in parentheses. m1 and m2 are serial correlation tests for differenced errors. * Rejection at the 5 percent.

Table C.9: Autorregresive Model of Earnings with Job Changes

	GMM1 Basic	GMM2 Basic	GMM2 By type of exit	GMM2 Same dynamics within and across	GMM2 No job-specific heterogeneity
y_{it-1}	0.060	0.026	0.018	0.149	0.272
	(0.047)	(0.043)	(0.044)	(0.021)	(0.048)
$d_{it-1}y_{it-1}$	0.133	0.153			0.096
	(0.051)	(0.049)			(0.086)
$d_{it-1}^{loss}y_{it-1}$			0.175		
			(0.083)		
$d_{it-1}^{quit}y_{it-1}$			0.161		
			(0.066)		
m1	-7.52*	-7.32*	-7.02*	-12.51*	-14.24*
m2	-0.42	-0.76	-0.76	0.35	1.72
Sargan test	-	292.38	291.25	297.59	301.09
(df)		(274)	(273)	(275)	(274)

Note: robust t-ratios in parentheses. m1 and m2 are serial correlation tests for differenced errors. * Rejection at the 5 percent.

Table C.10: Wage Variance Estimates

	Whole sample	Only stayers	Only movers	Only layoffs
$\left(\sigma_\mu^2 + \sigma_\phi^2\right)$	0.104	0.094	0.124	0.156
σ_μ^2	0.090	-	0.091	0.104
σ_ϕ^2	0.014	-	0.033	0.052
Obs.	19,069	9,064	10,005	2,014

Note: σ_μ^2 and σ_ϕ^2 are the variances of the individual and job effect. $\hat{\sigma}_\phi^2$ is obtained as the difference between $\left(\widehat{\sigma_\mu^2 + \sigma_\phi^2}\right)$ and $\hat{\sigma}_\mu^2$. Obs.: number of sample $\hat{u}_{it}$ available for calculation. I drop observations if consecutive changes for the same worker, and any sample covariance with less than 25 observations.

Appendix D

Tables of Chapter 4

Table D.1: Design 1: Static Probit for different values of T

			$T=8$	$T=12$
ML	$\hat{\beta}$	Mean	1.241	1.143
	$\hat{\beta}$	SD	0.133	0.087
BC-C	$\hat{\beta}$	Mean	1.118	1.052
	$\hat{\beta}$	SD	0.118	0.078
BC-E	$\hat{\beta}$	Mean	1.116	1.051
	$\hat{\beta}$	SD	0.117	0.078

Design: $y_{it} = 1\left[\eta_{i0} + \beta_0 x_{it} + \epsilon_{it} > 0\right]$;
$\epsilon_{it} \sim N(0,1)$; $x_{it} = 0.5x_{it-1} + u_{it}$;
$u_{it} \sim N(0,1)$; $x_{i0} \sim N(0,1)$; $\beta_0 = 1$;
$\eta_{i0} \sim N(0,1)$; N=100; 1,000 simulations;
SD: Sample Standard Deviation.

Table D.2: Design 2 with $\delta_0 = 1$: Static Probit for different values of T

			$T=6$	$T=8$	$T=10$	$T=12$
ML	$\hat{\beta}$	Mean	1.409	1.280	1.221	1.173
	$\hat{\beta}$	SD	0.078	0.049	0.041	0.041
	$\hat{\delta}$	Mean	1.386	1.242	1.182	1.145
	$\hat{\delta}$	SD	0.096	0.067	0.056	0.049
BC-C	$\hat{\beta}$	Mean	1.270	1.148	1.104	1.070
	$\hat{\beta}$	SD	0.072	0.044	0.036	0.032
	$\hat{\delta}$	Mean	1.239	1.113	1.069	1.045
	$\hat{\delta}$	SD	0.087	0.059	0.050	0.044
BC-E	$\hat{\beta}$	Mean	1.269	1.147	1.104	1.069
	$\hat{\beta}$	SD	0.072	0.044	0.036	0.032
	$\hat{\delta}$	Mean	1.236	1.111	1.067	1.043
	$\hat{\delta}$	SD	0.087	0.059	0.050	0.044

Design: $y_{it} = \mathbf{1}\left[\eta_{i0} + \beta_0 x_{it} + \delta_0 d_{it} + \epsilon_{it} > 0\right]$; $\epsilon_{it} \sim N(0,1)$;
$x_{it} \sim N(0,1)$; $d_{it} = \mathbf{1}[x_{it} + h_{it} > 0]$; $h_{it} \sim N(0,1)$;
$\eta_{i0} = \sqrt{T}\bar{x}_i + a_i, a_i \sim N(0,1)$; $\beta_0 = \delta_0 = 1$; N=1,000;
100 simulations; SD: Sample Standard Deviation.

Table D.3: Design 2 with $\delta_0 = 0.5$: Static Probit for different values of T

			$T=6$	$T=8$	$T=10$	$T=12$
ML	$\hat{\beta}$	Mean	1.360	1.248	1.196	1.157
	$\hat{\beta}$	SD	0.063	0.045	0.037	0.035
	$\hat{\delta}$	Mean	0.677	0.601	0.579	0.563
	$\hat{\delta}$	SD	0.082	0.059	0.053	0.043
BC-C	$\hat{\beta}$	Mean	1.220	1.123	1.087	1.061
	$\hat{\beta}$	SD	0.057	0.040	0.033	0.032
	$\hat{\delta}$	Mean	0.603	0.540	0.526	0.518
	$\hat{\delta}$	SD	0.073	0.052	0.048	0.040
BC-E	$\hat{\beta}$	Mean	1.218	1.122	1.087	1.060
	$\hat{\beta}$	SD	0.057	0.040	0.033	0.032
	$\hat{\delta}$	Mean	0.601	0.539	0.525	0.517
	$\hat{\delta}$	SD	0.073	0.052	0.048	0.040

Design: $y_{it} = \mathbf{1}\left[\eta_{i0} + \beta_0 x_{it} + \delta_0 d_{it} + \epsilon_{it} > 0\right]; \epsilon_{it} \sim N\left(0,1\right);$ $x_{it} \sim N\left(0,1\right); d_{it} = \mathbf{1}[x_{it} + h_{it} > 0]; h_{it} \sim N\left(0,1\right);$ $\beta_0 = 1; \delta_0 = 0.5; \eta_{i0} = \sqrt{T}\bar{x}_i + a_i,\ a_i \sim N\left(0,1\right)$;N=1,000; 100 simulations; SD: Sample Standard Deviation.

Table D.4: Design 3: Static Probit for different values of T

			$T=6$	$T=8$	$T=10$	$T=12$
ML	$\hat{\beta}$	Mean	1.294	1.190	1.147	1.117
	$\hat{\beta}$	SD	0.048	0.033	0.032	0.023
	$\hat{\delta}$	Mean	0.639	0.595	0.566	0.553
	$\hat{\delta}$	SD	0.064	0.048	0.045	0.034
BC-C	$\hat{\beta}$	Mean	1.155	1.072	1.047	1.031
	$\hat{\beta}$	SD	0.043	0.030	0.029	0.021
	$\hat{\delta}$	Mean	0.573	0.539	0.519	0.513
	$\hat{\delta}$	SD	0.058	0.044	0.041	0.031
BC-E	$\hat{\beta}$	Mean	1.155	1.072	1.047	1.031
	$\hat{\beta}$	SD	0.043	0.030	0.029	0.021
	$\hat{\delta}$	Mean	0.573	0.539	0.519	0.513
	$\hat{\delta}$	SD	0.058	0.044	0.041	0.031

Design: $y_{it} = \mathbf{1}\left[\eta_{i0} + \beta_0 x_{it} + \delta_0 d_{it} + \epsilon_{it} > 0\right]; \epsilon_{it} \sim N\left(0,1\right);$ $x_{it} \sim N\left(0,1\right); \eta_{i0} = 0, \forall i;$ $d_{it} = \mathbf{1}[x_{it} + h_{it} > 0]; h_{it} \sim N\left(0,1\right);$ $\beta_0 = 1; \delta_0 = 0.5;$ 100 simulations; N=1,000. SD: Sample Standard Deviation.

Table D.5: Design 4: Dynamic Probit for different values of T

			$T=8$	$T=12$
ML	$\hat{\beta}$	Mean	1.329	1.224
	$\hat{\beta}$	SD	0.149	0.110
	$\hat{\delta}$	Mean	0.056	0.210
	$\hat{\delta}$	SD	0.182	0.130
BC-C-Tri	$\hat{\beta}$	Mean	1.226	1.134
	$\hat{\beta}$	SD	0.139	0.101
	$\hat{\delta}$	Mean	0.287	0.374
	$\hat{\delta}$	SD	0.169	0.123
BC-E-Tri	$\hat{\beta}$	Mean	1.182	1.134
	$\hat{\beta}$	SD	0.221	0.101
	$\hat{\delta}$	Mean	0.272	0.372
	$\hat{\delta}$	SD	0.170	0.123
BC-C-Exp	$\hat{\beta}$	Mean	1.062	1.037
	$\hat{\beta}$	SD	0.106	0.087
	$\hat{\delta}$	Mean	0.390	0.442
	$\hat{\delta}$	SD	0.163	0.121
BC-E-Exp	$\hat{\beta}$	Mean	1.061	1.036
	$\hat{\beta}$	SD	0.106	0.087
	$\hat{\delta}$	Mean	0.402	0.448
	$\hat{\delta}$	SD	0.162	0.121

Design: $y_{it} = \mathbf{1}\left[\eta_{i0} + \beta_0 x_{it} + \delta_0 y_{it-1} + \epsilon_{it} > 0\right];$ $\epsilon_{it} \sim N(0,1); x_{it} = 0.5x_{it-1} + u_{it}; u_{it} \sim N(0,1);$ $x_{i0} \sim N(0,1); \eta_{i0} \sim N(x_{i0},1); \beta_0 = 1;$ $\delta_0 = 0.5;$ $y_{i0} = \mathbf{1}[\eta_i + \beta_0 x_{i0} + \epsilon_{i0} > 0]$; 200 simulations; N=100. Trimming =1; 200 samples. SD: Sample Standard Deviation.

Table D.6: Design 5: Dynamic Probit for different values of T

			$T=6$	$T=8$	$T=10$	$T=12$
ML	$\hat{\beta}$	Mean	1.267	1.183	1.135	1.106
	$\hat{\beta}$	SD	0.048	0.031	0.025	0.022
	$\hat{\delta}$	Mean	0.055	0.182	0.247	0.291
	$\hat{\delta}$	SD	0.055	0.038	0.034	0.032
BC-C-Tri	$\hat{\beta}$	Mean	1.187	1.102	1.061	1.040
	$\hat{\beta}$	SD	0.045	0.029	0.023	0.020
	$\hat{\delta}$	Mean	0.297	0.375	0.409	0.430
	$\hat{\delta}$	SD	0.051	0.035	0.032	0.030
BC-E-Tri	$\hat{\beta}$	Mean	1.188	1.103	1.062	1.040
	$\hat{\beta}$	SD	0.045	0.029	0.023	0.020
	$\hat{\delta}$	Mean	0.297	0.375	0.408	0.429
	$\hat{\delta}$	SD	0.051	0.034	0.032	0.030
BC-C-Exp	$\hat{\beta}$	Mean	1.021	1.013	1.006	1.002
	$\hat{\beta}$	SD	0.033	0.024	0.020	0.018
	$\hat{\delta}$	Mean	0.395	0.443	0.462	0.472
	$\hat{\delta}$	SD	0.047	0.034	0.032	0.030
BC-E-Exp	$\hat{\beta}$	Mean	1.021	1.012	1.006	1.002
	$\hat{\beta}$	SD	0.035	0.024	0.020	0.018
	$\hat{\delta}$	Mean	0.407	0.450	0.467	0.474
	$\hat{\delta}$	SD	0.047	0.034	0.032	0.030

Design: $y_{it} = \mathbf{1}\left[\eta_{i0} + \beta_0 x_{it} + \delta_0 y_{it-1} + \epsilon_{it} > 0\right]; \epsilon_{it} \sim N(0,1);$
$x_{it} \sim N(0,1); \eta_i = 0, \forall i; y_{i0} = \mathbf{1}[\eta_i + \beta_0 x_{i0} + \epsilon_{i0} > 0]; \beta_0 = 1;$
$\delta_0 = 0.5$; 100 simulations; N=1,000; Trimming =1; 250 samples;
SD: Sample Standard Deviation.

Table D.7: Design 6: Static Probit for different values of T

			$T=6$	$T=8$	$T=10$	$T=12$	$T=20$
ML	$\hat{\beta}$	Mean	1.569	1.416	1.322	1.264	1.144
	$\hat{\beta}$	SD	0.073	0.050	0.040	0.031	0.022
BC-C	$\hat{\beta}$	Mean	1.440	1.282	1.190	1.141	1.052
	$\hat{\beta}$	SD	0.069	0.045	0.036	0.028	0.019
BC-E	$\hat{\beta}$	Mean	1.436	1.279	1.187	1.139	1.051
	$\hat{\beta}$	SD	0.069	0.045	0.036	0.028	0.019

Design: $y_{it} = 1\left[\eta_{i0} + \beta_0 x_{it} + \delta_{i0} d_{it} + \epsilon_{it} > 0\right]; \epsilon_{it} \sim N(0,1); \beta_0$=1;
$x_{it} \sim N(0,1)$; $\eta_{i0} = 0, \delta_{i0} = 0.5, \forall i; d_{it} = 1[x_{it} + h_{it} > 0]; h_{it} \sim N(0,1)$;
100 simulations; N=1,000. SD: Sample Standard Deviation.

Table D.8: Design 7: Dynamic Probit for different values of T

			$T=6$	$T=8$	$T=10$	$T=12$	$T=20$
ML	$\hat{\beta}$	Mean	1.538	1.397	1.306	1.248	1.134
	$\hat{\beta}$	SD	0.062	0.046	0.035	0.031	0.019
BC-C-Tri	$\hat{\beta}$	Mean	1.458	1.295	1.198	1.143	1.052
	$\hat{\beta}$	SD	0.061	0.045	0.033	0.029	0.017
BC-E-Tri	$\hat{\beta}$	Mean	1.456	1.293	1.197	1.142	1.052
		SD	0.061	0.045	0.033	0.029	0.017

Design: $y_{it} = \mathbf{1}\left[\eta_{i0} + \beta_0 x_{it} + \delta_{i0} y_{it-1} + \epsilon_{it} > 0\right]; \epsilon_{it} \sim N(0,1); x_{it} \sim N(0,1)$; $\eta_{i0} = 0$,$\delta_{i0} = 0.5, \forall i$; y_{i0}= $1\left[\eta_{i0} + \beta_0 x_{i0} + \epsilon_{i0} > 0\right]; \beta_0 = 1$; 100 simulations; N=1,000. Trimming=1. SD: Sample Standard Deviation.

Appendix E

Figures

Figure E.1: Mean hourly wage by gender (pta 1992)

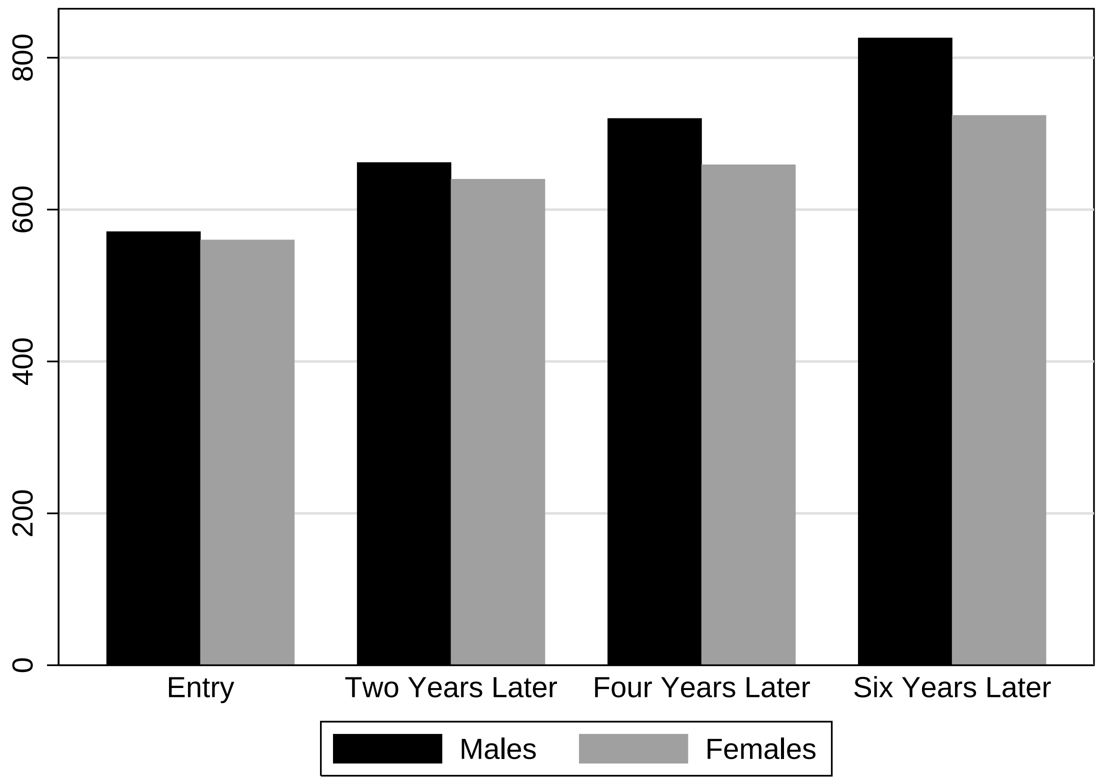


Figure E.2: Sample proportions by gender

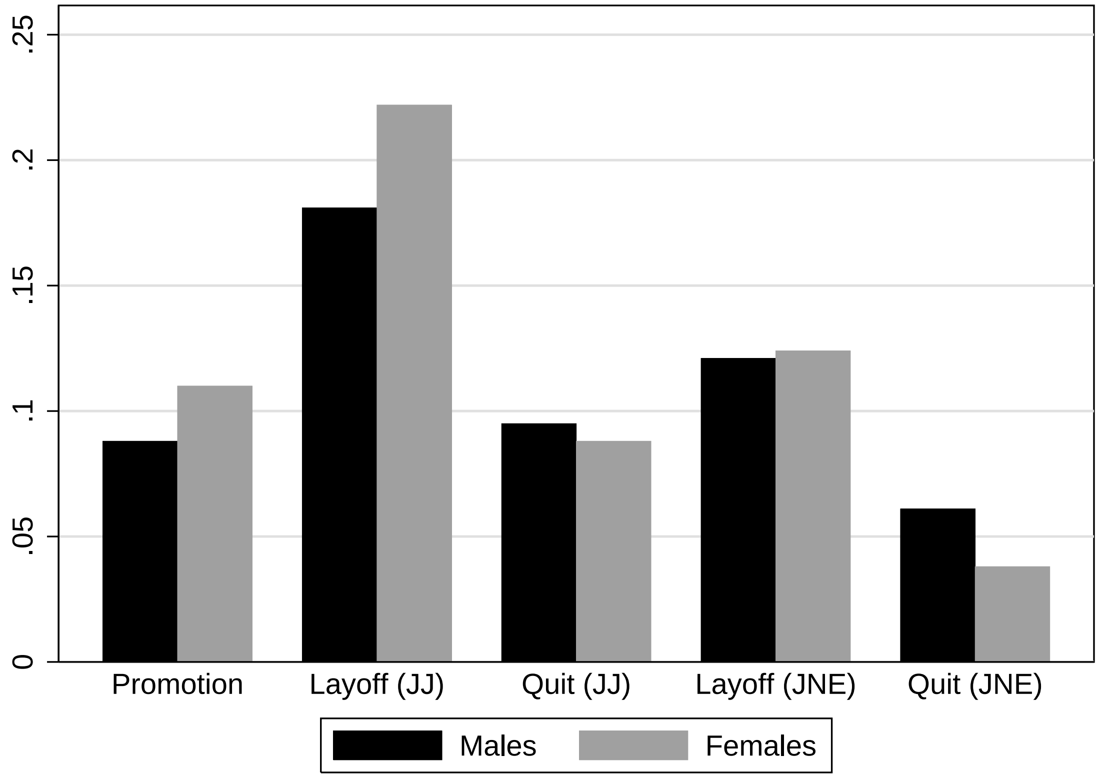


Figure E.3: The mean of log wages

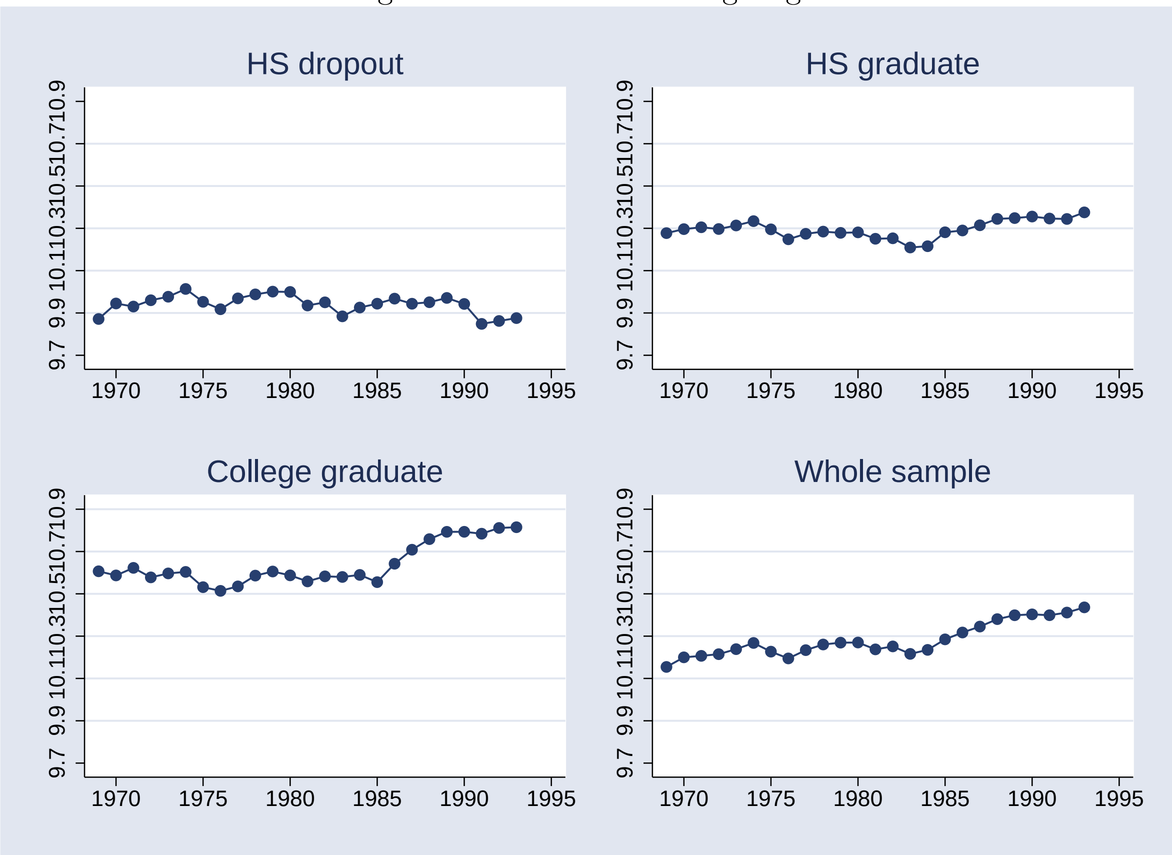


Figure E.4: The variance of log wages

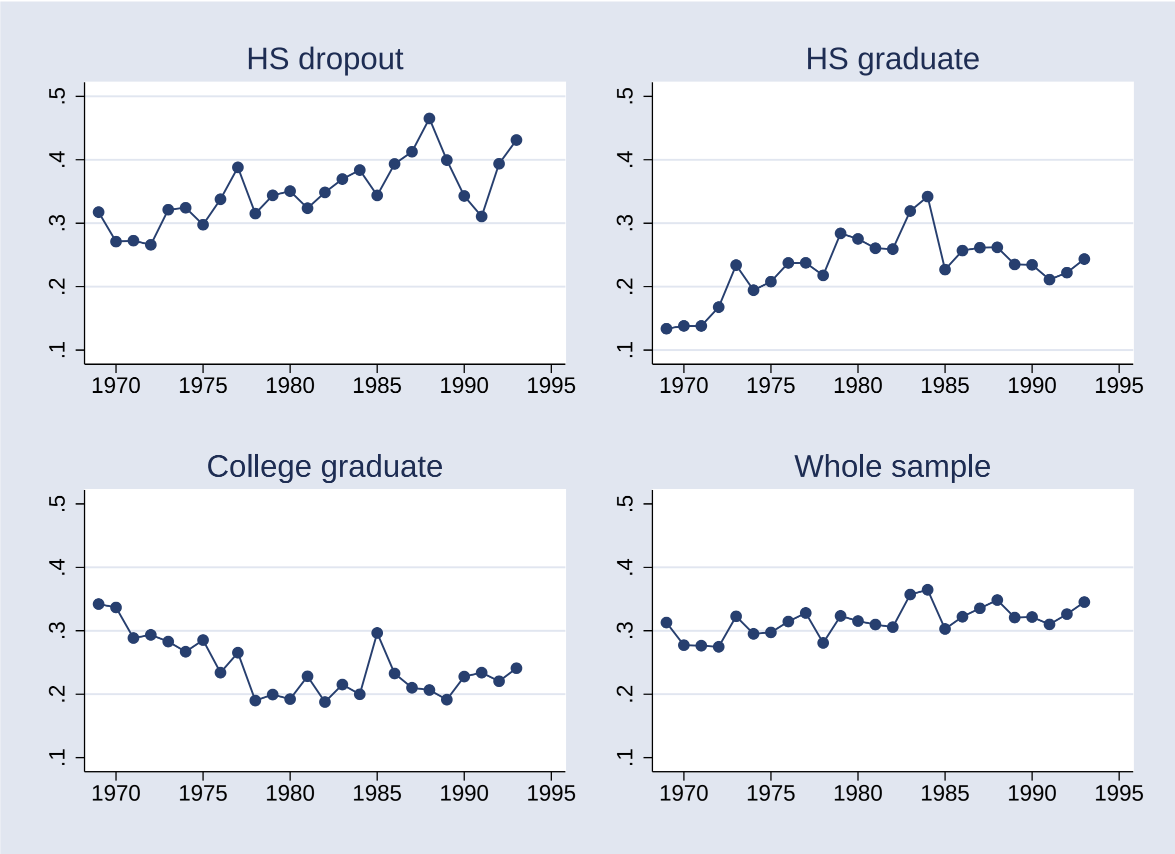


Figure E.5: Distribution of Residuals in First Differences

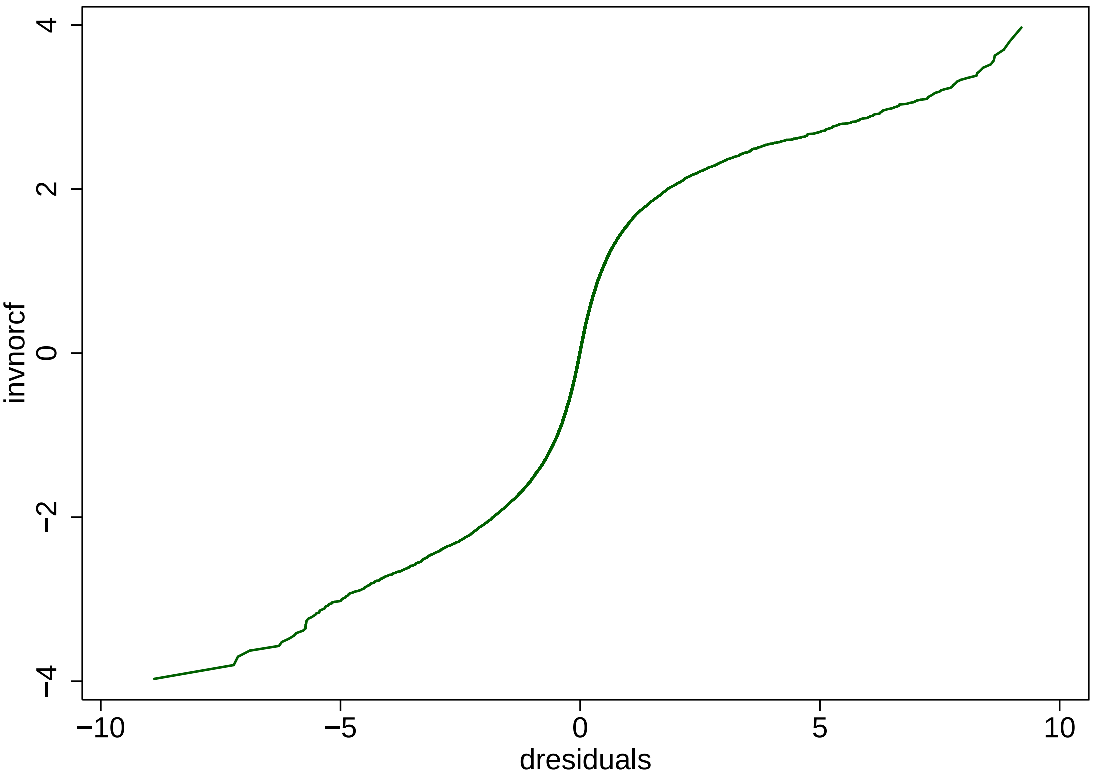


Figure E.6: Distribution of Standarized Residuals in First Differences

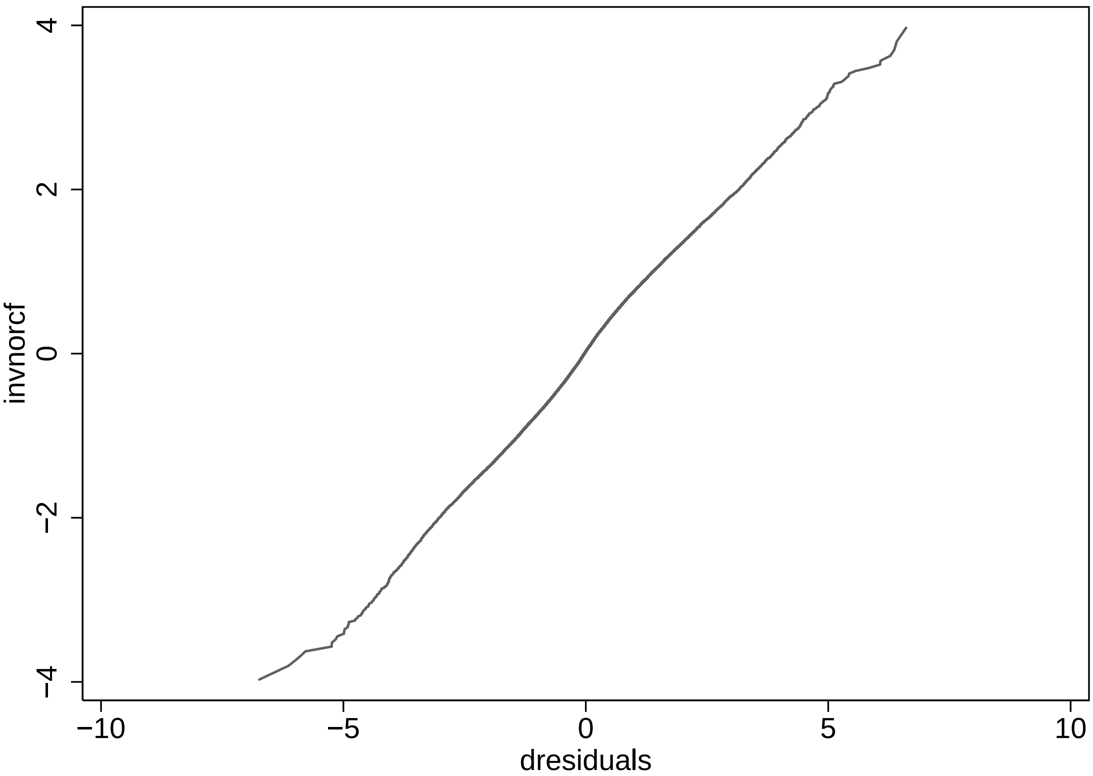


Figure E.7: Kernel densities of logwages and simulated logwages

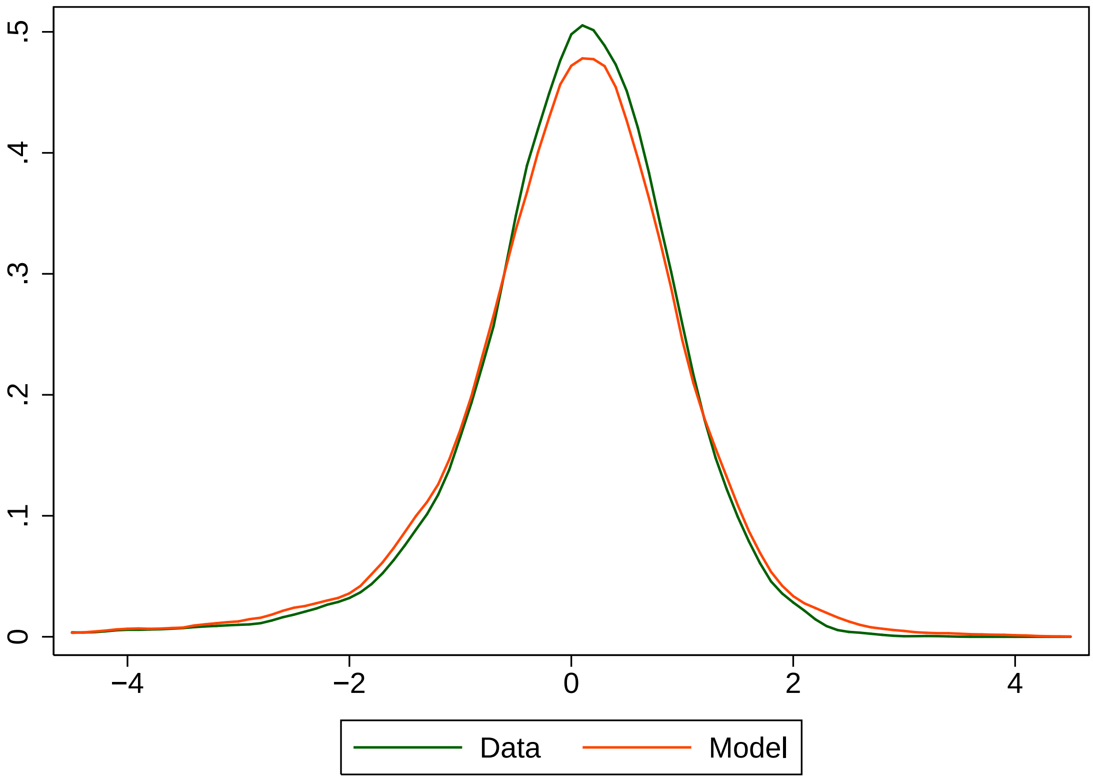


Figure E.8: Kernel density of individual means

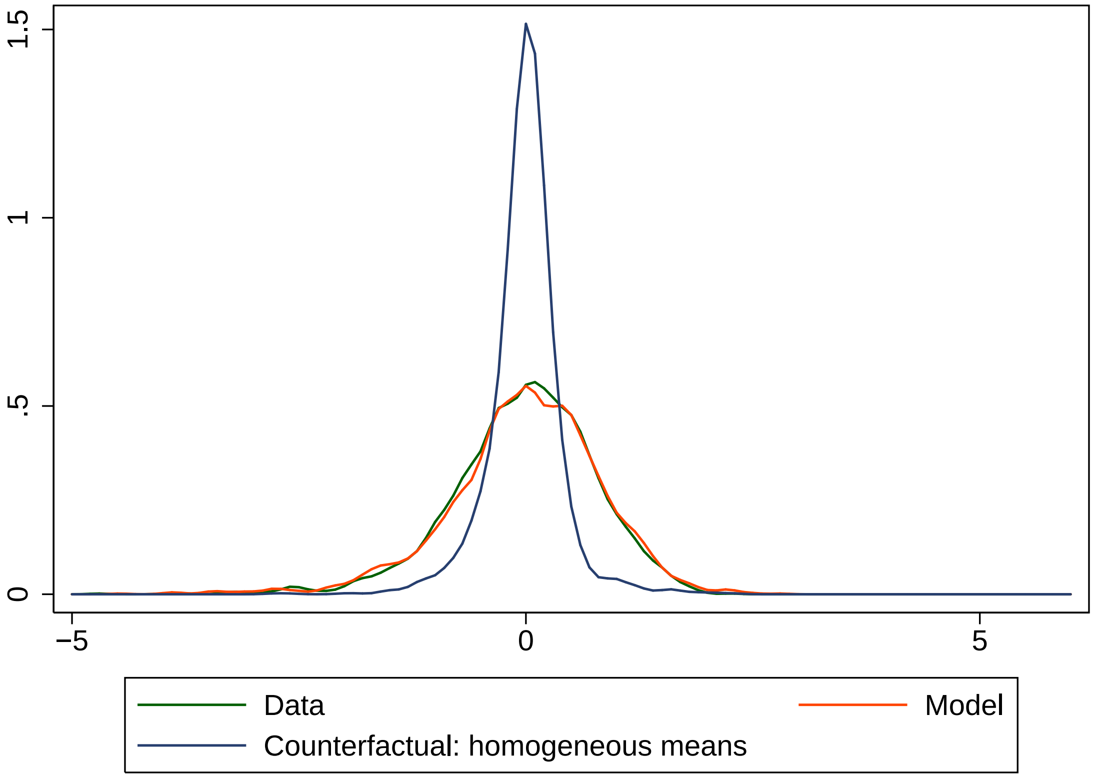


Figure E.9: Kernel density of individual logvariances

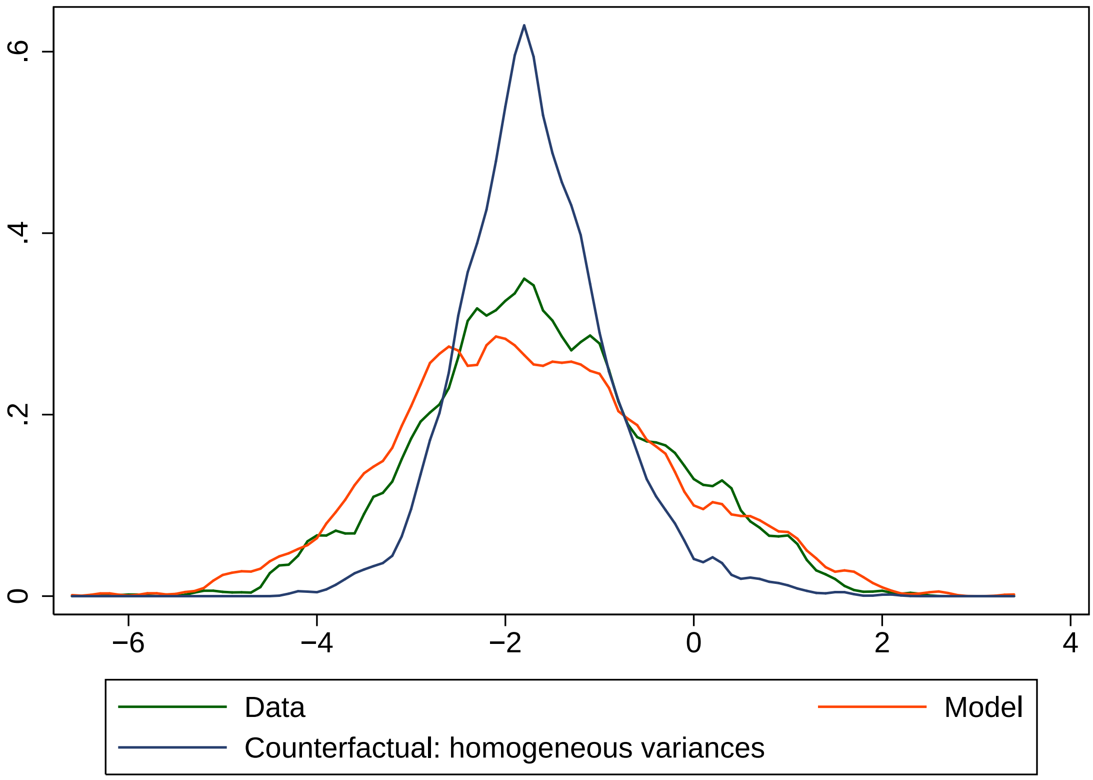


Figure E.10: Mean Elasticities

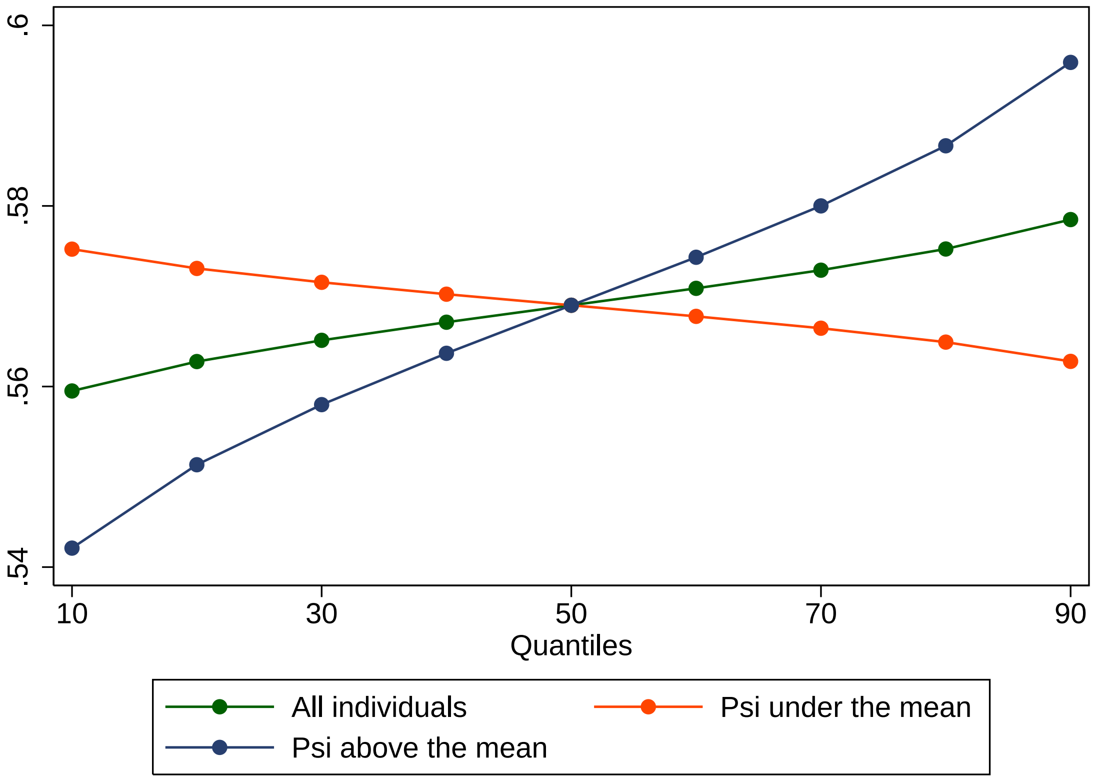


Figure E.11: Mean Marginal Effects

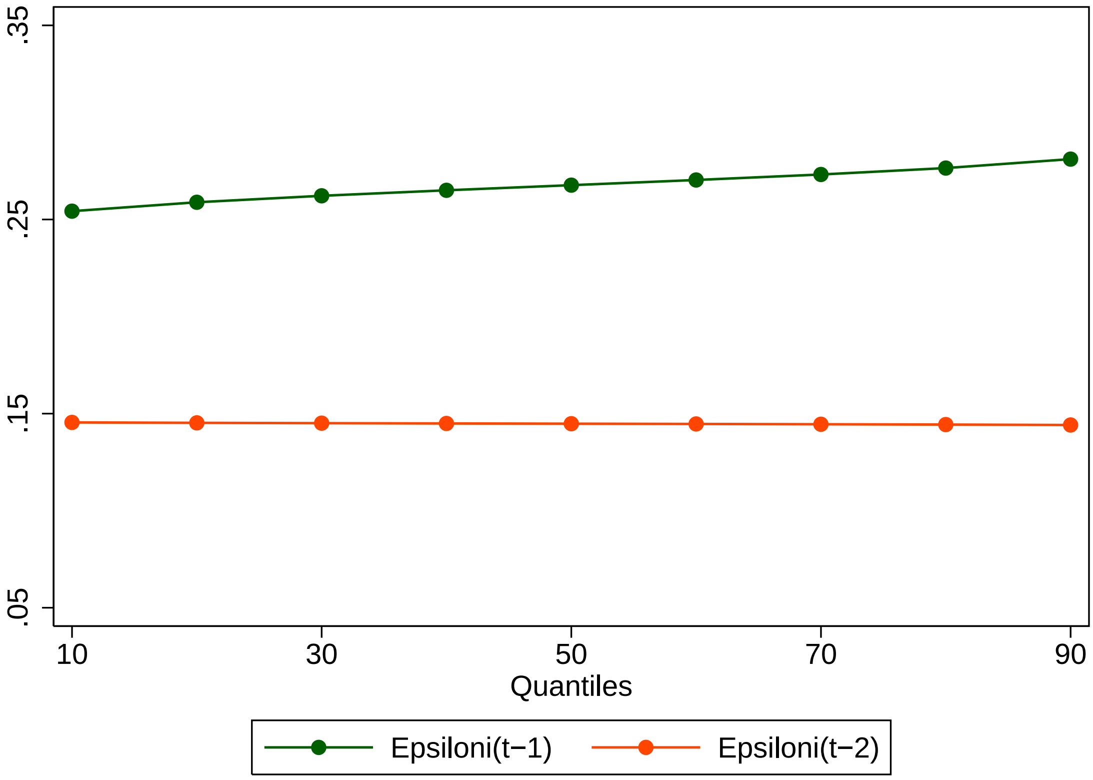


Figure E.12: Probability of Job Change

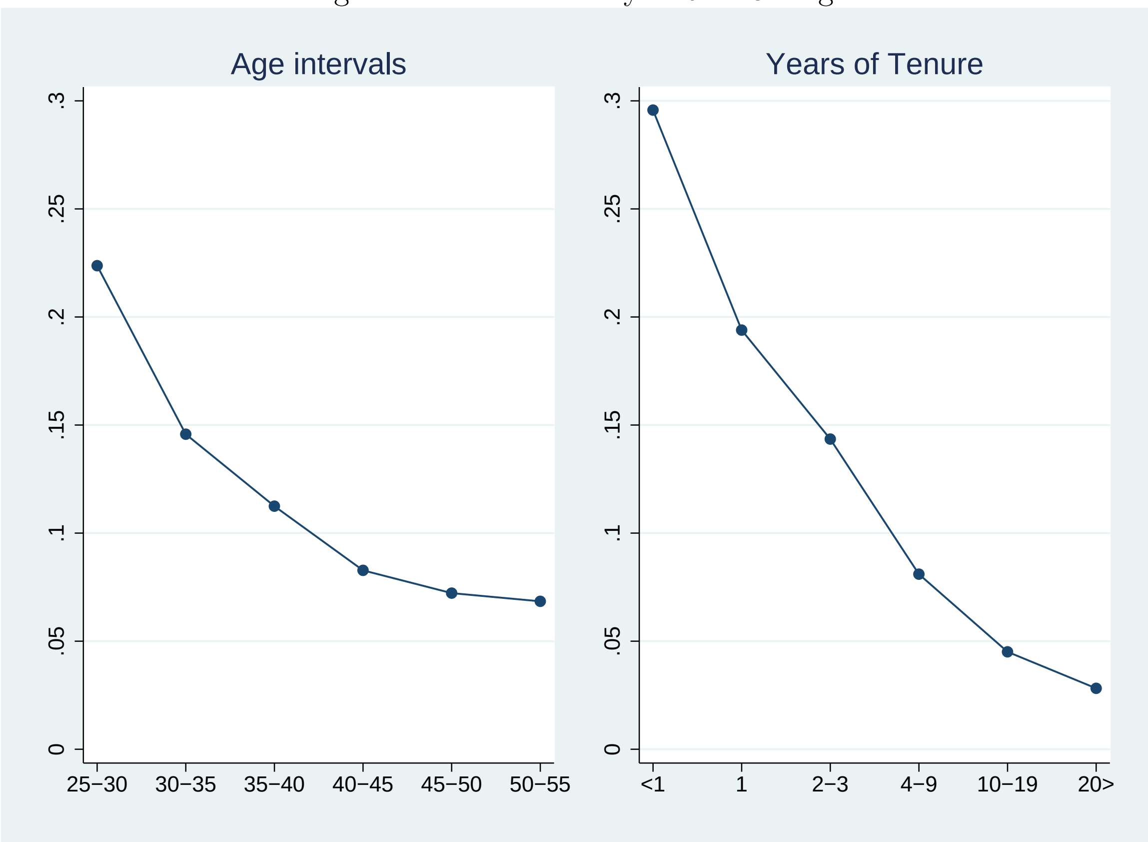
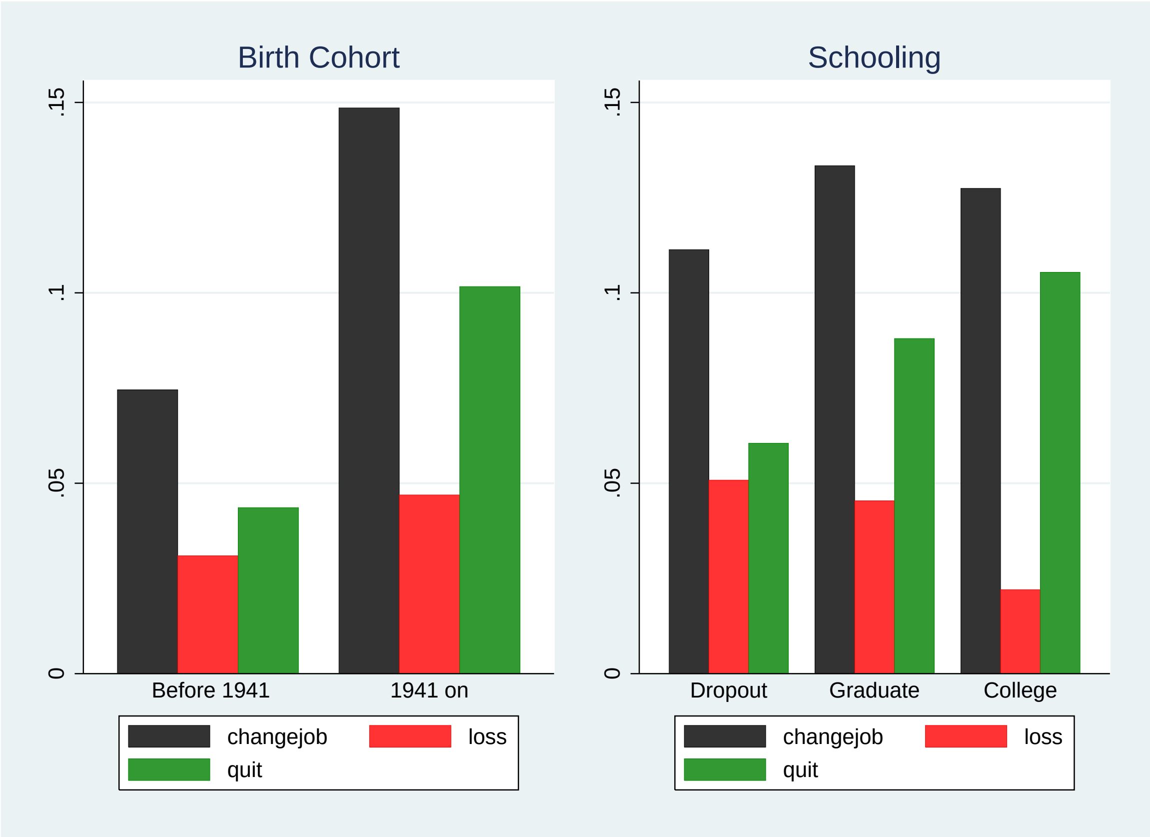


Figure E.13: Probability of Job Change by Exit Reason



Summary in Spanish / Resumen en Español

Esta tesis doctoral, titulada *Heterogeneity and Dynamics in Individual Wages and Labour Market Histories*, propone nuevos modelos y métodos de estimación para el análisis de los salarios y de los historiales laborales de los individuos desde una perspectiva dinámica. En el estudio de estos dos fenómenos, claves en el desarrollo profesional de los trabajadores, se utilizan datos de panel, esto es, observaciones de individuos que se repiten a lo largo del tiempo.

Por un lado, en los datos individuales observamos que los salarios evolucionan durante la vida laboral de los trabajadores. Dicha evolución será diferente según la fase del ciclo económico en la que se encuentre la economía así como en función de distintas características de los individuos, tanto características observables para el econometra (sexo, edad, nivel educativo, y muchas otras) como inobservables (como habilidad o suerte). Por otro lado, también se observa que el estatus laboral de los trabajadores cambia, pasando de la situación de empleado a no empleado, o de un puesto de trabajo a otro, ya sea voluntaria o involuntariamente. Estas transiciones de entrada y salida del mercado de trabajo, o de cambio de un empleo a otro, son frecuentes y, de nuevo, varían a nivel individual.

Por tanto, el punto de partida de esta tesis es la idea de que diferencias en los historiales laborales individuales pueden ayudar a entender la dinámica y heterogeneidad en la evolución de los salarios. Por ejemplo, en el caso de hombres y mujeres, esperaríamos que las diferencias por sexos en historiales laborales ayudasen a explicar una parte relevante de la brecha salarial residual. De hecho, en la literatura encontramos múltiples teorías que conectan la movilidad laboral con la existencia y persistencia de dicho diferencial por sexos a lo largo del tiempo. Así, se afirma que si la movilidad laboral de las mujeres está más restringida debido a variables como el lugar de residencia del marido o el cuidado de niños, las ganancias salariales predichas por los modelos de búsqueda y emparejamiento (Burdett, 1978; Jovanovic, 1979) serán menores (Keith y Williams, 1995). Argumentos similares podrían aplicarse al caso de individuos heterogéneos en otras dimensiones, bien sean observables (para el caso de la edad, Topel y Ward (1992) enfatizan la importancia que la movilidad tiene sobre el salario de los más jóvenes en Estados Unidos) o inobservables e, incluso, podrían extenderse al caso de heterogeneidad al nivel del emparejamiento entre individuo y puesto de trabajo (Postel - Vinay y Robin (2002) resaltan la relevancia de estos efectos de *match* o emparejamiento en un modelo con dinámica salarial dentro del mismo trabajo y entre distintos puestos).

En particular, esta tesis aborda cómo formular especificaciones que consideren diferentes niveles de heterogeneidad, tanto observable (capítulo 1) como inobservable (capítulos 2 y 3), individual (capítulo 2) y específica del puesto (capítulo 3), en modelos empíricos para la dinámica de la distribución de los salarios y las trayectorias laborales. El capítulo 4, por su parte, representa una aportación de carácter metodológico que puede ser de utilidad en múltiples aplicaciones económicas.

El primer capítulo estudia diferencias por sexos en el crecimiento salarial y la movilidad laboral de los jóvenes utilizando datos de la sección española del Panel de Hogares de la Unión Europea (1994-2001). Fijarse en el segmento de los jóvenes es relevante ya que ésta es la etapa de la vida laboral en la que se concentran los mayores incrementos salariales[1] En el capítulo se propone, en primer lugar, la construcción de una medida de experiencia

[1]Por ejemplo, Murphy y Welch (1990) encuentran que dos tercios del crecimiento salarial que un trabajador acumula a lo largo de su vida profesional se concentran en los primeros diez años.

que, al contrario que la experiencia potencial que se emplea normalmente, sí tiene en cuenta la existencia de interrupciones en la carrera profesional de los trabajadores. En segundo lugar, se plantea un análisis de los patrones de movilidad laboral de hombres y mujeres jóvenes.

De la comparación entre la medida propuesta - experiencia acumulada - y la medida usada habitualmente - experiencia potencial - resulta que los rendimientos salariales a la experiencia son más elevados con la medida acumulada y que esta diferencia es mayor para mujeres que para hombres. Este resultado apunta a la existencia de una penalización salarial por interrupciones más importante para las mujeres. En cuanto a los cambios de empleo, encuentro que las tasas de movilidad de los jóvenes son similares entre ambos sexos. Las diferencias surgen por el lado de los determinantes que resultan relevantes para cada grupo a la hora de cambiar de estatus, especialmente en caso de promoción o para las salidas al paro o la inactividad. Para los hombres, ostentar un puesto con responsabilidad o tener una familia a cargo resultan determinantes importantes de los cambios de empleos. Por el contrario, en el caso de las mujeres, importan características del puesto como el tipo de jornada o el tamaño de la empresa. Por último, además de la penalización de género por parar, en los datos también se observa que el crecimiento salarial de los hombres en los primeros años de la trayectoria laboral es mayor que para las mujeres, más aún en años en los que ocurren cambios de empleo. Las conclusiones de este primer capítulo están en línea con los resultados constatados en otros estudios recientes para datos de Estados Unidos[2] (Light y Ureta, 1992; Loprest, 1992), Italia (Del Bono y Vuri, 2006) y Finlandia (Napari, 2007)[3].

El segundo capítulo, parte central de la tesis, contribuye a la literatura sobre esti-

---

[2]En lo que sigue, EEUU.

[3]Por ejemplo, la evidencia empírica para EEUU (Loprest, 1992) indica que durante los primeros cuatro años en el mercado laboral, los hombres acumulan un crecimiento salarial del 36 por ciento frente a un 29 por ciento para las mujeres. Las cifras análogas en mi muestra serían un 26 y un 18 por ciento, respectivamente.

mación de procesos de renta proponiendo modelos que, además de considerar fuentes de heterogeneidad individual y dinámica en la media condicional de los ingresos (como en Lillard y Willis, 1978; MaCurdy, 1982; Abowd y Card, 1989, y muchos otros), incorporan fuentes adicionales de heterogeneidad y dinámica en la varianza condicional. Considerar las propiedades de la varianza de los ingresos laborales, además de la media, será importante para describir de una manera más rica los perfiles laborales de los individuos. Por ejemplo, un individuo que trata de predecir la evolución de su salario, para tomar decisiones hoy sobre cuánto ahorrar u otras cuestiones, estará interesado en anticipar no sólo el nivel de esos ingresos futuros sino también su varianza. Más aún, el comportamiento del individuo será diferente en el caso de que esa variabilidad tenga un componente de carácter permanente o se deba a un período de mayor inestabilidad que va a terminar. Por ello en este capítulo se propone un modelo dinámico de panel con efectos individuales tanto en la media como en la varianza condicional tipo ARCH (Autoregressive Conditional Heteroskedasticity).

La segunda contribución del capítulo es la de estimar este modelo en un contexto empírico específico, como son los datos de hombres americanos, cabeza de familia, del Panel Study of Income Dynamics (PSID) entre 1968 y 1993. Utilizar estos datos es interesante porque apenas existe evidencia sobre cómo evolucionan las volatilidades de los salarios individuales en un período de creciente desigualdad, esto es, de aumento en la varianza agregada (Juhn, Murphy, y Pierce, 1993, entre otros).

Desde el punto de vista metodológico, el capítulo aplica nuevos métodos de estimación basados en funciones de verosimilitud corregidas (DiCiccio y Stern, 1993; Severini, 1998a; Pace y Salvan, 2005; Arellano y Hahn, 2006b) adaptadas a un modelo dinámico para la varianza condicional con múltiples efectos fijos. Esta metodología permite reducir el sesgo de estimación en un contexto en el que el número de observaciones por individuo, T, es pequeño[4]. Dado que este tipo de correcciones no se han aplicado antes para modelos de panel, en primer lugar se evalúa su comportamiento en un contexto de muestras pequeñas mediante simulaciones de Monte Carlo. Los resultados de las simulaciones indican que el sesgo de estimación se corrige sustancialmente para diseños calibrados a los datos empleados en la aplicación empírica.

Los resultados de la parte empírica muestran que es importante tener en cuenta la presencia tanto de heterogeneidad individual inobservable como de dinámica en la varianza condicional de los salarios y que ésta última está relacionada con la movilidad laboral. En una muestra parecida, Meghir y Pistafferi (2004) también encuentran fuerte evidencia tanto de heterogeneidad individual como de dinámica significativa en las varianzas. Estos autores consideran un modelo de panel ARCH para la dinámica de los ingresos laborales y calculan condiciones de momentos para estimarlo. Su método depende críticamente del supuesto de especificación lineal para la varianza. Incluso en ese caso, reconocen que en la práctica no pueden tener estimaciones consistentes debido a un problema de instrumentos débiles. El orden del sesgo del estimador que implementan es $1/T$, frente a un $1/T^2$ en el caso del estimador corregido de sesgo. Esta diferencia es muy importante como muestran las simulaciones en la comparación con el estimador de máxima verosimilitud, también de orden $1/T$. En mi modelo propongo una especificación exponencial que implica una varianza condicional siempre no negativa independientemente del valor de los parámetros (Nelson, 1992), pero lo que resulta interesante del método de estimación empleado en

[4]Los métodos de estimación de modelos de panel no lineales con corrección de sesgo constituyen una reciente línea de investigación en microeconometría (para una recopilación véase Arellano y Hahn, 2006a). Además de las correcciones de sesgo para la verosimilitud concentrada del tipo consideradas aquí, también hay métodos automáticos basados en simulación (Hahn y Newey, 2004), correcciones basadas en ortogonalización (Cox y Reid, 1987; Lancaster, 2002) y sus extensiones (Woutersen, 2002; Arellano, 2003), correcciones de sesgo analíticas de los estimadores (Hahn y Newey, 2004; Hahn y Kuersteiner, 2004) y correcciones de sesgo de las ecuaciones de momentos (Carro, 2006; Fernández - Val, 2005). La ventaja fundamental de las correcciones en la verosimilitud es que la expresión para el sesgo es más simple que en el estimador o en el *score*, especialmente con múltiples efectos fijos, además de que contar con un modelo completamente especificado nos permite calcular contrafactuales e, incluso, evaluar el ajuste del modelo a los datos.

este capítulo es que no depende de la formulación particular del modelo[5]. Un resultado adicional del capítulo es que el modelo explica la no normalidad que se observa en los datos de salarios en logaritmos.

Por último, el capítulo ilustra las implicaciones del modelo de salarios sobre el crecimiento del consumo en el marco de un modelo sencillo de ahorro por motivo precaución (Browning y Lusardi, 1996). La principal conclusión es que un aumento del riesgo a nivel individual induce una reducción significativa en el consumo actual y este efecto es más importante para el grupo de los menos educados, algo significativo para los individuos con educación secundaria y no significativo para los universitarios. Una posible interpretación de este resultado sería que estos últimos individuos son los que disponen de mayores posibilidades de aseguramiento[6].

Directamente conectado con el capítulo 2, el tercer capítulo desarrolla un modelo que considera explícitamente los cambios de empleo en la dinámica de los salarios y en la configuración de la heterogeneidad. Se trata de un modelo de datos de panel dinámico, del tipo denominado de componentes de error, que puede utilizarse para examinar el impacto que los cambios de empleo tienen sobre la dinámica y los distintos componentes de la varianza de los salarios individuales. En particular, la especificación propuesta permite que el parámetro que mide la dinámica en los salarios dentro de un mismo empleo sea diferente del que corresponde a la dinámica en años en los que hay cambio. Del mismo modo, el patrón de heterogeneidad inobservable se hace más rico permitiendo que además de un componente individual permanente en todos los períodos, exista también otro componente específico del puesto y que - por tanto - variará de un empleo al siguiente.

Dentro de las innumerables referencias que, dentro de la economía laboral, se han centrado en el estudio de los salarios, podemos distinguir dos vertientes. Por un lado,

[5]De hecho podría utilizarse sin modificaciones importantes en una especificación como la de Meghir y Pistaferri (2004).

[6]Interpretación coherente con los resultados de Blundell, Pistaferri y Preston (2005).

muchos artículos se han fijado en el estudio de los determinantes de los salarios. Algunos de estos artículos, basándose en teorías del capital humano (Becker, 1975), examinan el impacto de la experiencia general sobre los salarios. Muchos otros se fijan en el efecto del capital humano de carácter específico, basándose en teorías de búsqueda y emparejamiento (Burdett, 1978; Jovanovic 1979) o el puro *learning by doing* (Rosen, 1972), y estiman rendimientos salariales a la experiencia y la antigüedad (Altonji y Shakotko, 1987; Topel, 1991; Topel y Ward, 1992; Neal, 1995; Altonji y Williams, 1997; Dustmann y Meghir, 2005; entre otros) tratando de controlar la endogeneidad de la antigüedad con métodos diversos. Por otro lado, existe una literatura - relacionada con la anterior - pero que se ha preocupado más bien de modelizar y estimar las propiedades de serie temporal de los ingresos laborales, y que en su mayoría ha ignorado la distinción entre dinámica dentro de un empleo y entre empleos. El modelo propuesto en el capítulo 3 se encuadraría en la segunda corriente pero tomando de la primera la atención en la movilidad laboral y la preocupación por el carácter endógeno de estos cambios. Así, en el modelo propuesto se controla por la potencial endogeneidad de la movilidad introduciendo explícitamente los dos componentes de heterogeneidad inobservable, tanto individual como específica del puesto, y a la hora de establecer las condiciones de momentos que nos dan la identificación de los parámetros y que permiten su estimación, se tiene en cuenta que los cambios de empleo pueden estar correlacionados con esos componentes y también con shocks pasados, esto es, con la configuración de la historial del individuo en el pasado.

En la aplicación empírica utilizo, de nuevo, datos procedentes del PSID. El estudio se centra en los denominados cambios *job-to-job*, esto es, en transiciones de un trabajo a otro. Además la información contenida en el PSID permite establecer la distinción entre cambios voluntarios e involuntarios (como en caso de despido o cierre de la empresa). Los principales resultados son los que siguen. En cuanto a la dinámica, encuentro que una vez que controlamos por efectos de individuo y de puesto la dinámica dentro de un mismo empleo deja de ser significativa, mientras que es positiva aunque no muy elevada en años en los que se producen cambios. Establecer la distinción entre cambio voluntario o involuntario no resulta relevante para la dinámica pero sí en el caso de los componentes de la varianza, que estimo combinando covarianzas muestrales de sección cruzada entre observaciones para individuos en el mismo empleo y observaciones del mismo trabajador en distintos puestos. De este modo, para los individuos que cambian, encuentro que la estimación de la varianza del componente de heterogeneidad en el puesto representa un tercio de la varianza en la heterogeneidad netamente individual. Si considero una submuestra en la que cada cambio de empleo ocurre sólo involuntariamente (despidos o cierre de la empresa), obtengo que la varianza debido a la heterogeneidad en puestos aumenta hasta representar la mitad de la varianza entre individuos.

El siguiente paso inmediato en mi agenda de investigación sería la comparación de los resultados obtenidos para EEUU en los capítulos 2 y 3 con los correspondientes para diferentes países europeos; así como la extensión del modelo endogeneizando la decisión de participación, lo que haría posible la inclusión de las mujeres en el análisis y la consideración de transiciones al estado de paro o inactividad.

Por último, el capítulo 4 supone una contribución fundamentalmente de carácter técnico, relacionada con el cálculo computacional en la práctica de las correcciones de sesgo del tipo considerado en el capítulo 2. En este cuarto capítulo se considera la estimación de modelos de panel no lineales que incluyen múltiples efectos fijos individuales. La estimación de estos modelos en la práctica es difícil por dos razones. En primer lugar, en un modelo de este tipo puede haber hasta cientos o incluso miles de coeficientes individuales para estimar, además de un número relativamente pequeño de parámetros comunes. El segundo problema, conocido como el *Incidental Parameters Problem* (Neyman y Scott, 1948), surge porque las estimaciones de los efectos fijos cuando la dimensión temporal es corta serán muy imprecisas, lo que contaminará las estimaciones de los parámetros de interés debido a la no linealidad del modelo.

Una solución computacional muy utilizada en el caso lineal consiste en obtener primero las estimaciones de máxima verosimilitud (MV) de los parámetros comunes a partir de una regresión de los datos transformados en desviaciones respecto a las medias individuales y, a continuación, recuperar estimaciones MV de los efectos, uno por uno, promediando los residuos. Existe también una simplificación computacional similar para el algoritmo de Newton - Raphson para la estimación de modelos no lineales con efectos fijos que explota la estructura del hesiano diagonal en bloques[7]. El primer objetivo del capítulo es mostrar como usar un algoritmo de este tipo en un modelo no lineal con múltiples efectos fijos y, el segundo objetivo, discutir su aplicación a las funciones de verosimilitud corregidas. Los resultados se ilustran mediante un ejercicio de simulaciones de Monte Carlo para varios diseños.

Este último capítulo representa otra interesante línea de futura investigación, ya que aún son necesarios más resultados sobre cómo estas correcciones de sesgo funcionan en la práctica en diferentes modelos económicos y en más micropaneles y bases de datos de interés para la econometría aplicada, así como resultados sobre las propiedades teóricas que pueden ayudarnos a la hora de seleccionar entre los diferentes métodos de corrección disponibles.

[7]Esta modificación para modelos no lineales con un efecto fijo escalar se discute en Hall (1978), Chamberlain (1980), y Greene (2004).

Summary in Galician / Resumo en Galego

Esta tese de doutoramento, titulada *Heterogeneity and Dynamics in Individual Wages and Labour Market Histories*, propón novos modelos e métodos de estimación para a análise dos salarios e dos historiais de traballo dos individuos desde unha perspectiva dinámica. No estudo destes dous fenómenos, claves no desenvolvemento profesional dos traballadores, utilízanse datos de panel, isto é, observacións de individuos que se repiten ó longo do tempo.

Por unha banda, nos datos individuais observamos que os salarios evolucionan durante a vida laboral dos traballadores. Devandita evolución será diferente segundo a fase do ciclo económico na que se atope a economía así como en función de distintas características dos individuos, tanto características observables para o económetra (sexo, idade, nivel educativo, e moitas outras) como inobservables (como habilidade ou sorte). Por outra banda, tamén se observa que o estatus laboral dos traballadores cambia, pasando da situación de empregado a non empregado, ou dun posto de traballo a outro, xa sexa voluntaria ou involuntariamente. Estas transicións de entrada e saída do mercado de traballo, ou de cambio dun emprego a outro, son frecuentes e, de novo, varían a nivel individual.

Xa que logo, o punto de partida desta tese é a idea de que diferencias nos historiais de traballo individuais poden axudar a entender a dinámica e heteroxeneidade na evolución dos salarios. Por exemplo, no caso de homes e mulleres, esperariamos que as diferencias por sexos nos historiais axudasen a explicar unha parte relevante da brecha salarial residual. De feito, na literatura atopamos múltiples teorías que conectan a mobilidade laboral coa existencia e persistencia de devandito diferencial por sexos ó longo do tempo. Así, afírmase que se a mobilidad laboral das mulleres está máis restrinxida debido a variables como o lugar de residencia do marido ou o coidado dos nenos, as ganancias salariais preditas polos modelos de procura e emparellamento (Burdett, 1978; Jovanovic, 1979) serán menores (Keith e Williams, 1995). Argumentos similares poderían aplicarse ó caso de individuos heteroxéneos noutras dimensións, ben sexan observables (para o caso da idade, Topel e Ward (1992) enfatizan a importancia que a mobilidade ten sobre o salario dos máis novos en Estados Unidos) ou inobservables e, ata, poderían estenderse ó caso da heteroxeneidade ó nivel do emparellamento entre individuo e posto de traballo (Postel - Vinay e Robin (2002) resaltan a relevancia destes efectos de match ou emparellamento nun modelo con dinámica salarial dentro do mesmo traballo e entre distintos postos).

En particular, esta tese aborda como formular especifiacións que consideren diferentes niveis de heteroxeneidade, tanto observable (capítulo 1) como inobservable (capítulos 2 e 3), individual (capítulo 2) e específica do posto (capítulo 3), en modelos empíricos para a dinámica da distribución dos salarios e as traxectorias laborais. O capítulo 4, pola súa banda, representa unha achega de carácter metodolóxica que pode ser de utilidade en múltiples aplicacións económicas.

O primeiro capítulo estuda diferencias por sexos no crecemento salarial e a mobilidade laboral dos novos empregando datos da sección española do Panel de Fogares da Unión Europea (1994-2001). Fixarse no segmento dos traballadores novos é relevante xa que esta é a etapa da vida laboral na que se concentran os maiores incrementos salariais[8]. No capítulo proponse, en primeiro lugar, a construcción dunha medida de experiencia que,

[8]Por exemplo, Murphy e Welch (1990) atopan que dous terzos do crecemento salarial que un traballador acumula ó longo da súa vida profesional concéntranse nos primeiros dez anos.

ó contrario que a experiencia potencial que se emprega normalmente, si teña en conta a existencia de interrupcións na carreira profesional dos traballadores. En segundo lugar, suscítase unha análise dos patróns de mobilidade laboral de homes e mulleres xóvenes.

Da comparación entre a medida proposta - experiencia acumulada - e a medida usada habitualmente - experiencia potencial - resulta que os rendementos salariais á experiencia son máis elevados coa medida acumulada e que esta diferencia é maior para mulleres que para homes. Este resultado apunta á existencia dunha penalización salarial por interrupcións máis importante para as mulleres. En canto ós cambios de emprego, encontro que as taxas de mobilidade dos traballadores novos son similares entre ambos sexos. As diferencias xorden polo lado dos determinantes que resultan relevantes para cada grupo á hora de cambiar de estatus, especialmente en caso de promoción ou para as saídas ó paro ou á inactividade. Para os homes, ostentar un posto con responsabilidade ou ter unha familia a cargo resultan determinantes importantes dos cambios de empregos. Pola contra, no caso das mulleres, importan características do posto como o tipo de xornada ou o tamaño da empresa. Para rematar, ademais da penalización de xénero por parar, nos datos tamén se observa que o crecemento salarial dos homes nos primeiros anos da traxectoria laboral é maior que para as mulleres, máis aínda en anos nos que ocorren cambios de emprego. As conclusións deste primeiro capítulo están en liña cos resultados constatados noutros estudos recentes para datos de Estados Unidos[9] (Light e Ureta, 1992; Loprest, 1992), Italia (Del Bono e Vuri, 2006) e Finlandia (Napari, 2007)[10].

O segundo capítulo, parte central da tese, contribúe á literatura sobre estimación de procesos de renda propoñendo modelos que, ademais de considerar fontes de heteroxeneidade individual e dinámica na media condicional dos ingresos (como en Lillard e Willis,

[9]En adiante, EEUU.

[10]Por exemplo, a evidencia empírica para EEUU (Loprest, 1992) indica que durante os primeiros catro anos no mercado laboral, os homes acumulan un crecemento salarial do 36 por cento fronte a un 29 por cento para as mulleres. As cifras análogas na miña mostra serían un 26 e un 18 por cento, respectivamente.

1978; MaCurdy, 1982; Abowd e Card, 1982, e moitos outros), incorporan fontes adicionais de heteroxeneidade e dinámica na varianza condicional. Considerar as propiedades da varianza dos ingresos laborais, ademais da media, será importante para describir dun xeito máis rico os perfís laborais dos individuos. Por exemplo, un individuo que trata de predecir a evolución do seu salario, para tomar decisións hoxe sobre canto aforrar ou outras cuestións, estará interesado en anticipar non só o nivel deses ingresos futuros senón tamén a súa varianza. Máis aínda, o comportamento do individuo será diferente no caso de que esa variabilidade teña un compoñente de carácter permanente ou se deba a un período de maior inestabilidade que vai terminar. Por iso neste capítulo proponse un modelo dinámico de panel con efectos individuais tanto na media como na varianza condicional tipo ARCH (Autoregressive Conditional Heteroskedasticity).

A segunda contribución do capítulo é a de estimar este modelo nun contexto empírico específico, como son os datos de homes americanos, cabezas de familia, do Panel Study of Income Dynamics (PSID) entre 1968 e 1993. Utilizar estes datos é interesante porque apenas existe evidencia sobre como evolucionan as volatilidades dos salarios individuais nun período de crecente desigualdade, isto é, de aumento na varianza agregada (Juhn, Murphy, e Pierce, 1993, entre outros).

Desde o punto de vista da metodoloxía, o capítulo aplica novos métodos de estimación baseados en funcións de verosimilitude corrixidas (DiCiccio e Stern, 1993; Severini, 1998a; Pace e Salvan, 2005; Arellano e Hahn, 2006b) adaptadas a un modelo dinámico para a varianza condicional con múltiples efectos fixos. Esta metodoloxía permite reducir o sesgo de estimación nun contexto no que o número de observacións por individuo, T, é pequeno[11]. Dado que este tipo de correccións non se aplicaron antes para modelos de

[11] Os métodos de estimación de modelos de panel non lineais con corrección de sesgo constitúen unha recente liña de investigación en microeconometría (para unha recompilación ver Arellano e Hahn, 2006a). Ademais das correccións de sesgo para a verosimilitude concentrada do tipo consideradas aquí, tamén hai métodos automáticos baseados en simulación (Hahn e Newey, 2004), correccións baseadas en ortogonalización (Cox e Reid, 1987; Lancaster, 2002) e as súas extensións (Woutersen, 2002; Arellano, 2003), correccións de sesgo analíticas dos estimadores (Hahn e Newey, 2004; Hahn e Kuersteiner, 2004) e

panel, en primeiro lugar avalíase o seu comportamento nun contexto de mostras pequenas mediante simulacións de Monte Carlo. Os resultados das simulacións indican que o sesgo de estimación corríxese substancialmente para deseños calibrados ós datos empregados na aplicación empírica.

Os resultados da parte empírica mostran que é importante ter en conta a presencia tanto da heteroxeneidade individual inobservable como da dinámica na varianza condicional dos salarios e que esta última está relacionada coa mobilidade laboral. Nunha mostra parecida, Meghir e Pistafferi (2004) tamén atopan forte evidencia tanto de heteroxeneidade individual como de dinámica significativa nas varianzas. Estes autores consideran un modelo de panel ARCH para a dinámica dos ingresos laborais e calculan condicións de momentos para estimalo. O seu método depende criticamente do suposto de especificación linear para a varianza. Ata nese caso, recoñecen que na práctica non poden ter estimacións consistentes debido a un problema de instrumentos débiles. A orde do sesgo do estimador que implementan é $1/T$, fronte a un $1/T^2$ no caso do estimador corrixido de sesgo. Esta diferencia é moi importante como mostran as simulacións na comparación co estimador de máxima verosimilitude, tamén de orde $1/T$. No meu modelo propoño unha especificación exponencial que implica unha varianza condicional sempre non negativa independentemente do valor dos parámetros (Nelson, 1992), pero o que resulta interesante do método de estimación empregado neste capítulo é que non depende da formulación particular do modelo[12]. Un resultado adicional do capítulo é que o modelo explica a non normalidade que se observa nos datos de salarios en logaritmos.

Para rematar, o capítulo ilustra as implicacións do modelo de salarios sobre o crece-

correccións de sesgo das ecuacións de momentos (Carro, 2006; Fernández - Val, 2005). A vantaxe fundamental das correccións na verosimilitude é que a expresión para o sesgo é máis simple que no estimador ou no *score*, especialmente con múltiples efectos fixos, ademais de que contar cun modelo completamente especificado permítenos calcular contrafactuales e ata avaliar o axuste do modelo ós datos.

[12]De feito podería utilizarse sen modificacións importantes nunha especificación como a de Meghir e Pistaferri (2004).

mento do consumo no marco dun modelo sinxelo de aforro por motivo precaución (Browning e Lusardi, 1996). A principal conclusión é que un aumento do risco a nivel individual induce unha reducción significativa no consumo actual e este efecto é máis importante para o grupo dos menos educados, algo significativo para os individuos con educación secundaria e non significativo para os universitarios. Unha posible interpretación deste resultado sería que estes últimos individuos son os que dispoñen de maiores posibilidades de aseguramento[13].

Directamente conectado co capítulo 2, o terceiro capítulo desenvolve un modelo que considera explícitamente os cambios de emprego na dinámica dos salarios e na configuración da heteroxeneidade inobservable. Trátase dun modelo de datos de panel dinámico, do tipo denominado de compoñentes de erro, que pode utilizarse para examinar o impacto que os cambios de emprego teñen sobre a dinámica e os distintos compoñentes da varianza dos salarios individuais. En particular, o modelo permite que o parámetro que mide a dinámica nos salarios dentro dun mesmo emprego sexa diferente do que corresponde á dinámica en anos nos que hai cambio. Do mesmo xeito, o patrón de heteroxeneidade inobservable faise máis xeral permitindo que ademais dun compoñente individual permanente en todos os períodos, exista tamén outro compoñente específico do posto e que - polo tanto - variará dun emprego ó seguinte.

Dentro das innumerables referencias que, dentro da economía laboral, centráronse no estudo dos salarios, podemos distinguir dúas vertentes. Por unha banda, moitos artigos fixáronse no estudo dos determinantes dos salarios. Algúns destes artigos, baseándose en teorías do capital humano (Becker, 1975), examinan o impacto da experiencia xeral sobre os salarios. Moitos outros se fixan no efecto do capital humano de carácter específico, baseándose en teorías de procura e emparellamento (Burdett, 1978; Jovanovic 1979) ou o puro *learning by doing* (Rosen, 1972), e estiman rendementos salariais á experiencia e

[13]Interpretación coherente cos resultados de Blundell, Pistaferri e Preston (2005).

á antigüidade (Altonji e Shakotko, 1987; Topel, 1991; Topel e Ward, 1992; Neal, 1995; Altonji e Williams, 1997; Dustmann e Meghir, 2005; entre outros) tratando de controlar pola endoxeneidade da antigüidade con métodos diversos. Por outra banda, existe unha literatura - relacionada coa anterior - pero que se preocupou máis ben de modelizar e estimar as propiedades de serie temporal dos ingresos laborais, pero que na súa maioría ignorou a distinción entre dinámica dentro dun emprego e entre empregos. O modelo proposto neste capítulo 3 encadraríase na segunda corrente pero tomando da primeira a atención pola mobilidade laboral e a preocupación polo posible carácter endóxeno destes cambios. Así, no modelo proposto contrólase pola potencial endoxeneidade da mobilidade introducindo explicitamente os dous compoñentes de heteroxeneidade inobservable, tanto individual como específica do posto, e á hora de establecer as condicións de momentos que nos dan a identificación dos parámetros e que permiten a súa estimación, tense en conta que os cambios de emprego poden estar correlacionados con eses compoñentes e tamén con shocks pasados, isto é, coa configuración do historial do individuo no pasado.

Na aplicación empírica utilizo, de novo, datos procedentes do PSID. O estudo céntrase nos denominados cambios *job-to-job*, isto é, en transicións dun traballo a outro. Ademais a información contida no PSID permite establecer a distinción entre cambios voluntarios e involuntarios (como en caso de despedimento ou pechadura da empresa). Os principais resultados son os que seguen. En canto á dinámica, encontro que unha vez que controlamos por efectos de individuo e de posto, a dinámica dentro dun mesmo emprego deixa de ser significativa, mentres que é positiva aínda que non moi elevada en anos nos que se producen cambios. Establecer a distinción entre cambio voluntario ou involuntario non resulta relevante para a dinámica, pero si no caso dos compoñentes da varianza, que estimo combinando covarianzas muestrais de sección cruzada entre observacións para individuos no mesmo emprego e observacións do mesmo traballador en distintos postos. Deste xeito, para os individuos que cambian, encontro que a estimación da varianza do compoñente de heteroxeneidade no posto representa un terzo da varianza na heteroxeneidade netamente individual. Se considero unha submostra na que cada cambio de emprego ocorre só involuntariamente (despedimentos ou pechadura da empresa), obteño que a varianza debida á heteroxeneidade en postos aumenta ata representar a metade da varianza entre individuos.

O seguinte paso inmediato na miña axenda de investigación sería a comparación dos resultados obtidos para os EEUU nos capítulos 2 e 3 cos correspondentes para diferentes países europeos; así como a extensión do modelo endoxeneizando a decisión de participación, o que faría posible a inclusión das mulleres na análise e a consideración de transicións ó estado de paro ou inactividade.

Para rematar, o capítulo 4 supón unha contribución de carácter técnico, relacionada co cálculo computacional na práctica das correccións de sesgo do tipo considerado no capítulo 2. Neste cuarto capítulo se considera a estimación de modelos de panel non lineais que inclúen múltiples efectos fixos individuais. A estimación destes modelos na práctica é difícil por dúas razóns. En primeiro lugar, nun modelo deste tipo poder haber centos ou ata miles de coeficientes individuais para estimar, ademais dun número relativamente pequeno de parámetros comúns. O segundo problema, coñecido como o *Incidental Parameters Problem* (Neyman e Scott, 1948), xurde porque as estimacións dos efectos fixos cando a dimensión temporal é curta serán moi imprecisas, o que contaminará as estimacións dos parámetros comúns debido á non linearidade do modelo.

Unha solución computacional moi empregada no caso linear consiste en obter primeiro as estimacións de máxima verosimilitude (MV) dos parámetros comúns a partir dunha regresión dos datos transformados en desviacións respecto das medias individuais e, a continuación, recuperar estimacións MV dos efectos, un por un, promediando os residuos. Existe tamén unha simplificación computacional similar para o algoritmo de Newton - Raphson para a estimación de modelos non lineais con efectos fixos que explota a estrutura do hesiano diagonal en bloques[14]. O primeiro obxectivo do capítulo é mostrar como usar un algoritmo deste tipo nun modelo non linear con múltiples efectos fixos e, o segundo, discutir a súa aplicación ás funcións de verosimilitude corrixidas. Os resultados ilústranse mediante un exercicio de simulacións de Monte Carlo para varios deseños.

Este último capítulo representa outra interesante liña de futura investigación, xa que aínda son necesarios máis resultados sobre como estas correccións de sesgo funcionan na práctica para diferentes modelos económicos e en máis micropaneles e bases de datos de interese para a econometría aplicada, así como resultados sobre as propiedades teóricas que poden axudarnos á hora de seleccionar entre os diferentes métodos de corrección do sesgo dispoñibles.

[14]Esta modificación para modelos non lineais cun efecto fixo escalar discútese en Hall (1978), Chamberlain (1980), e Greene (2004).