Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Yue Li

The Covariance Structure of Male Wages in the United Kingdom

MSc Thesis 2012-021

The Covariance Structure of Male Wages in the United Kingdom

Yue Li

March 2012

Master's thesis in Econometrics, Operation Research and Actuarial Studies
Supervisor: Prof. R.J.M. Alessie
Co-assessor: Prof. R.H. Koning
Author: Yue Li (Student No. s1659987)

Abstract

This paper uses the New Earnings Survey (NES) panel dataset from 1975 to 2001 to analyse the variance-covariance structure of individual gross hourly wage in the United Kingdom. Special emphasis is given to the evidence for the validity of the Restricted Income Profile (RIP) hypothesis versus the Heterogeneous Income Profile (HIP) hypothesis. These two processes are typically used to model the individual earnings process. Applying minimum distance estimation to the error components model, I find evidence in favor of the RIP hypothesis. In the UK individuals are subject to large and persistent income shocks while facing similar life-cycle earnings profiles. This finding is in contrast with the previous studies using the US data, which conclude that the HIP hypothesis holds and individuals are subject to modest persistent income shocks while facing individual-specific earnings profiles. Under the RIP process, this paper also discusses the persistence of the transitory component and its accumulated effect on cross-sectional variance. Furthermore, the analysis of the age-variance profile implies that the cross-sectional variance increases with age in a concave fashion, and this finding is in line with the previous empirical literature using the UK data.

Keywords: Covariance structure; Income stochastic process; Minimum distance estimation; Optimization

Preface

This thesis is the final project of my master degree in Econometrics, Operations Research and Actuarial Studies, specialisation in Econometrics. At the beginning I picked up this topic from my supervisor Prof. Alessie's research theme of Netspar, and this thesis was originally designed to study the individual income process in the Netherlands. Unfortunately, due to the deficiencies of the Duthch Inkomens Panel Onderzoek (IPO) dataset, I have to change the dataset to the NES and modify the contents of this thesis.

I would like to express my gratitude to Prof. Alessie, not only for his supervision on this thesis, but also for his guidance on my study throughout these years and his recommendation letters and information for my PhD application. I would also like to thank my second supervisor Prof. Koning for his time on this thesis, as well as his teaching and kind explanations to my questions during his lectures.

Second, I would like to thank all my friends and classmates. As being an international student, you make my life more colorful. Special thanks give to Sofia Vounatsou and Dirk Sackman, my life would be much more difficult without your help.

Finally, I want to thank to my parents. I alway know that I should work harder to express my appreciations for your thoughtful supports during my life.

Contents

List of Tables

List of Figures

1 Introduction

Over the last forty or so years, there has been modest or strong increases in wage inequality in the US (Gottschalk and Moffitt, 1994; Haider, 2001; Shin and Solon, 2011), the UK (Dickens, 2000; Ramos, 2003; Kalwij and Alessie, 2007) and Canada (Baker and Solon, 2003; Beach, Finnie, and Gray, 2010). This increase in cross-sectional wage inequality is either caused by an increase in permanent shocks, which are due to permanent characteristics such as skilled-based technical change, or comes from transitory shocks that have temporary and limited persistent effects on individual's earnings. The nature of permanent and transitory shocks makes it important for policies aiming at reducing wage inequality to clearly distinguish these two types of shocks. If the wage inequality is predominately caused by transitory shocks, the wage inequality would be reduced to a lower level in the next period and the policies are less urgent and less necessary. However, the policies will only have a short time effect in case that permanent shock is the main reason for wage inequality, unless the policy makers analyse the underlying causes of permanent shocks.

The individual earnings process has been applied to a range of economics and finance models. A direct implication is in life-cycle consumption behavior. For example, Attanasio and Weber (2010) show that consumption and saving respond in different ways to permanent and transitory shocks of income, given that transitory shocks are more easily to be smoothed than permanent shocks. Permanent and transitory shocks also have implications for social welfare theory. Shorrocks (1978) and Gottschalk and Spolaore (2000) assume that social welfare depends on the possibility for individuals to change their rank in the income distribution over their life-cycles. They argue that an increase in permanent shocks leads to social-welfare-detracting, since it moves individuals further apart from the income distribution and makes the individual's income rank less possible to change. In contrast, because an increase in transitory shocks mixes up the income distribution, it brings social-welfare-improving. Starting with the individual income process, Storesletten et al. (2004) use cross-sectional variation between cohorts of similar ages to identify the cyclical idiosyncratic labor-market shocks. Moreover, the dynamic earnings process is the basis of the determination of wealth inequality (Huggett, 1996), asset prices (Constantinides and Duffie, 1996) and the welfare cost of business cycles (Lucas, 2003).

The analyses on the individual earnings process have been widely documented over the past decades. In general, most of the studies are based on the US data. Early studies, exemplified by Lillard and Willis (1978), MaCurdy (1982), Gottschalk (1982) and Abowd and Card (1989), fit various simple models to the sample variance-covariance matrices of individual incomes. One limitation of these studies is that they have not considered possible changes in the variance-covariance structure of earnings. Later on, taking into account the changes in parameters over time, Gottschalk and Moffitt (1995) use the Panel Study of Income Dynamics (PSID) data of the US to disentangle permanent shocks from transitory shocks in wage inequality and estimate the percentage of both components that explains the rise in wage inequality between 1969 and 1987. The disentanglement of the permanent and transitory shocks has further implications for the studies in wage mobility, since a large permanent component or a high persistence of transitory shocks implies a lower level of wage mobility. Gittleman and Joyce (1994) do not find evidence that there is a change in short run mobility in the US by using Current Population Survey from 1967 to 1991. However, Buchinsky and Hunt (1999) find evidence of falling mobility for young workers in the US for the same time period by using National Longitudinal Survey of Young from 1979 to 1991.

The UK data is also frequently utilized by substantial recent studies. By using the British Family Expenditure Survey and information on household total expenditure, Blundell and Preston (1995) find that transitory income inequality has been increased in the late 1980s and early 1990s. The British Household Panel Study is firstly used by Ramos (2003), who concludes that earnings persistence has fallen and mobility has increased over 1990s. Dickens (2000) and Kalwij and Alessie (2007) use the New Earnings Survey (NES) from 1975-1995 and 1975-2001 respectively. Dickens (2000) finds that the variance of the permanent and transitory components have increased in the early and late 1980s respectively, each explaining about half of the rise in the income inequality. By allowing transitory wage inequality to depend on age in a non-parametric way, Kalwij and Alessie (2007) control for the age and cohort effects and provide a more extended and flexible model of Dickens (2000) and Ramos (2003). They identify a continuing increase in wage inequality up to 2001, and attribute this increase mainly to a strong rise in transitory wage inequality.

When disentangling the income process into permanent and transitory components and modeling the former component, it is typical for current studies to use either a Restricted Income Profile (RIP) process or Heterogeneous Income Profile (HIP) process. In the first process, individuals are assumed to be subject to large and persistent income shocks while facing similar life-cycle profiles. Studies that use this process include Dickens (2000), Ramos (2003) and Kalwij and Alessie (2007). In the HIP process, individuals are assumed to be subject to income shocks with modest persistence and face individual-specific earnings profiles. A random growth model is usually used to model this process (Haider, 2001). One particular topic of the individual earnings process has focused on examining the evidence for the validity of the HIP and the RIP process. Using Panel Study of Income Dynamics (PSID), Guvenen (2009) shows that the HIP process is suitable for the individual income process in the US and the estimated persistence of shocks is lower than that of the RIP process.

An important and interesting question is whether or not there exists any empirical evidence that leads to support the HIP process in other countries except the US. In this paper, using the UK New Earnings Survey (NES) panel dataset from 1975 to 2001, I reassess the methodology and results in Guvenen (2009). In contrast with his findings and conclusions, my results suggest that there exists no empirical evidence to support the HIP process. Therefore, my study provides statistical support to the studies that assume and use the RIP process to model the earnings process in the UK, exemplified by Dickens (2000), Ramos (2003) and Kalwij and Alessie (2007).

The rest of the paper is organized as follows. In Section 2, I discuss the New Earnings Survey (NES) panel dataset and provide descriptive statistics. In Section 3, I introduce the methodology and estimation procedure. In Section 4, I present and discuss the empirical results and in Section 5 I conclude.

2 The New Earnings Survey

2.1 Introduction to the Survey

The New Earnings Survey (NES) is an annual administrative dataset of the earnings of individuals in employment in Great Britain, with the aim to collect information about the levels, distribution, composition of earnings and other deductions. It is designed and collected by National Statistics to cover all categories of occupations in labor market. The NES contains one percent sample of all employees who are members of Pay As you Earn (PAYE) income tax schemes, and is randomly drawn over age sixteen based on individual's unique personal National Insurance (NI) number. The information in this survey is collected from PAYE records around one month before the beginning of the financial year and the survey pay-period. Due to the data collection procedure, the coverage of part-time employees is not comprehensive. The earnings below the income tax threshold are not covered in this survey, therefore it has excluded most of the information on part-time jobs, self-employed workers and a small proportion of women, students and young people. This survey ends in 2003 and has been replaced by the Annual Survey of Hours and Earnings (ASHE). [1]

The NES provides plenty of advantages and convenience for researchers. In total it covers 29 years and this substantial time period is extended compared with other administrative panel datasets of personal earnings. For instance, there are only 15 years observation in German Socio-Economic Panel (GSOEP). Furthermore, the questionnaire remains fairly consistent during these years and the same NI number is used as the basis for each year's sample, providing the possibility to generate the annual cross-section panel. On average NES records the information on half a million individuals, and this number is much larger than that in PSID of the US (22000 in year 1968 and 70000 individuals in all). In addition, the majority of employers (around 75%) are contacted through the Inland Revenue tax register using the PAYE income tax schemes records and employers

[1] Annual Survey of Hours and Earnings (ASHE) provides a more representative sample of employees and more comprehensive coverage of all employees, especially for low-paid employees (Bird, 2004).

are obliged to complete the questionnaire on the basis of payroll records for the employee, therefore the response rate can be highly insured and measurement errors are significantly reduced. The sample frame also makes sure that an individual will not be excluded in the future years if he/she is not observed due to some reasons in any year. For more detailed dicusssion on individuals entrance and exit of panel please refer to Kalwij and Alessie (2007).

Inevitably, there are also a number of drawbacks in the NES. Because the NES is mostly used as a cross-sections survey, National Statistics did not pay enough attention to the panel aspect of the survey. Potential selection problem may arise from sample attrition in the NES. Those who are before retirement or in employment may be absent from the survey in some years. This is caused by the following two main reasons: the individual's earnings falls below the threshold that is required to pay income tax; or the individual changes jobs during the one month time between locating this individual and when employer receives the request from the survey. Most empirical studies only select the sample of adult male wage and impose restrictions on minimum wage, so the samples are not likely to be affected by the income tax thresholds. Due to the second reason, it is possible that the NES undersamples the individuals with high rate of job turnover. Although the New Earnings Survey contains the information on individual's working hours, industry, occupation, place of work, sex, age, and distribution of earnings, there is a lack of information on individual's personal characteristics, such as education level, health condition, family situation and working experience. Therefore, we cannot control the potential problem that the variance-covariance structure of earnings may depend on those factors.

2.2 Data Selection and Variable Definition

The individual earnings data are drawn from 1975 to 2001 of the NES. The original unbalanced survey contains half a million individuals and 4.4 million total observations. In this study, I strictly follow the data selection procedure used in Kalwij and Alessie (2007). At first 3.8% of the observations which have a gender inconsistency are dropped. Since there are potential sample selection issues concerning with female and young individuals wage distribution, I only select male employees (57%) who are aged between 21 and 59 (inclusive, 86% of male employees) and have provided information on the hours of work (85.8% of male aged between 21 and 59). At last, for the sake of robustness of estimation results and filting out extreme observations, the individuals who have more than one job and the observations that are at the top and bottom 0.1% of the wage distribution are excluded. The above selection procedure results in approximately 66000 men in each year, 219495 distinct individuals and 1.77 million total number of observations from year 1975 to 2001.

The variable of individual earnings used in this paper is the hourly wage rate, which is defined as the standard gross earnings including overtime payment divided by the number of working hours. Wage rates are deflated to 2001 pounds by the Retail Price Index.

2.3 Comparison with PSID Data

It is informative to briefly describe how Guvenen (2009) selects the data and variables in his study. Guvenen (2009) uses the US labor earnings data which are drawn from the first 26 waves of PSID from 1968 to 1993. The sample used in his empirical research satisfies the following conditions: the observations are male head of households between ages of 20 and 64; the individuals should be observed for at least 20 years (not necessarily consecutive) from 1968 to 1993; employee has a positive income and the working hours ranging from 520 to 5110 in a year; individuals's income does not belong to poverty (SEO) sample in 1968; extremely high/low observations are filtered out.

The most important difference between the data selection procedure of Guvenen (2009) and mine is that Guvenen (2009) requires individuals to be observed for at least twenty years. This restriction makes his sample to contain less heterogeneous individuals. Furthermore, Kalwij and Alessie (2007) and I do not impose any restrictions for minimum and maximum working hours. Other criteria are very similar. Concerning with the data size, the number of observations in this paper is much larger than that of Guvenen (2009). On average there are around 66000 observations in each year after the selection procedure, while this number is only around 1000 in Guvenen (2009).

2.4 Descriptive Findings

The descriptive statistics of male wages for year 1975 to 2001 is reported in Table 1. As discussed in Section 2.2, after the selection procedure there are still around 60000 to 70000 observations in each year, providing an adequate number of observations to calculate the sample variance-covariance matrix of individual wages. The average (log-) wage rate keeps on increasing from 1975 to 2001. The cross-sectional variance of the log-wage and the ratio of the 90th and 10th percentile of the wage distribution can be regarded as the measures for cross-sectional wage inequality. As discussed in Dickens (2000) and Kalwij and Alessie (2007), the massive increase in cross-sectional wage inequality starts around 1978 and stops in 1995, remaining fairly stable afterwards.

Year	Average age	Average wage rate	Standard deviation	90th/10th deciles	Average log-wage	Variance of log-wage	Number of observations
1975	39.757	7.776	3.279	2.371	1.983	0.124	66355
1976	39.676	7.831	3.409	2.438	1.986	0.131	70576
1977	39.572	7.235	3.028	2.368	1.912	0.122	70625
1978	39.664	7.568	3.232	2.429	1.954	0.128	70041
1979	39.618	7.845	3.343	2.441	1.989	0.131	69888
1980	39.595	7.908	3.448	2.506	1.994	0.137	69472
1981	39.310	8.173	3.774	2.657	2.018	0.152	68025
1982	39.060	8.227	3.862	2.713	2.022	0.158	67318
1983	39.005	8.608	4.114	2.787	2.064	0.165	66017
1984	38.910	8.804	4.384	2.850	2.081	0.172	64762
1985	38.819	8.796	4.382	2.874	2.080	0.174	62464
1986	38.655	9.211	4.719	2.935	2.122	0.181	64507
1987	38.521	9.376	5.068	3.037	2.131	0.196	64313
1988	38.377	9.823	5.502	3.141	2.172	0.206	67181
1989	38.455	9.984	5.729	3.171	2.184	0.212	66456
1990	38.488	10.000	5.749	3.205	2.184	0.215	66254
1991	38.445	10.300	5.990	3.304	2.209	0.225	65308
1992	38.418	10.517	6.134	3.384	2.228	0.230	62544
1993	38.249	10.801	6.462	3.436	2.249	0.241	61170
1994	38.301	10.751	6.673	3.489	2.239	0.250	61897
1995	38.524	10.784	6.985	3.642	2.232	0.270	66229
1996	38.665	10.922	7.049	3.637	2.245	0.268	65378
1997	38.862	11.135	7.184	3.610	2.265	0.264	62325
1998	39.055	11.217	7.413	3.646	2.269	0.270	65164
1999	39.191	11.541	7.698	3.682	2.296	0.269	64703
2000	39.522	11.691	7.808	3.665	2.310	0.267	62433
2001	39.679	12.079	8.325	3.779	2.334	0.281	63539

Note: The individual wage used in this paper is the hourly wage rate, which is defined as the standard gross hourly earnings including overtime payment. Wage rates are deflated to 2001 pounds by the Retail Price Index.

Table 1: Descriptive statistics for all years

Figure 1 depicts the relation between age and cross-sectional variance of log-wage for selected birth cohorts. The general information emerging from this figure is the positive correlation between age and cross-sectional variance. Given a fixed age, it always holds that the variance of younger cohort is larger than the variance of older cohort. It can be concluded that those individuals in younger cohort face a larger wage inequality.

Figure 2 plots how the variance-covariance changes over time. For all orders, both variance and covariance increase with time in a similar time pattern. In general, the covariance of each lag decreases with order at a decreasing rate, but remains at a positive level in all twenty-five years. This phenomena implies the existence of a permanent component in the variance-covariance of log-wages. It is also interesting to notice the parallelism and the nearly constant distance among the lines, which may be explained by a possible stable error structure in the individual earnings process. So the theoretical structural function should take into account the above features when fitting the sample variance-covariance of log-wage.

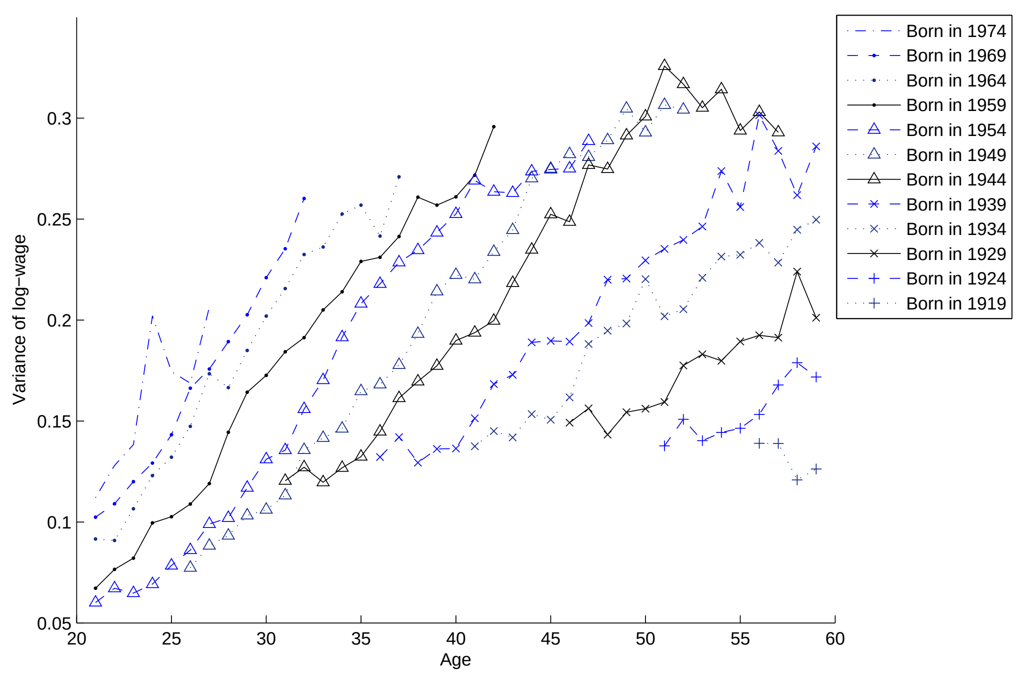

Figure 1: Lifetime variance of log-wage for selected birth cohort.

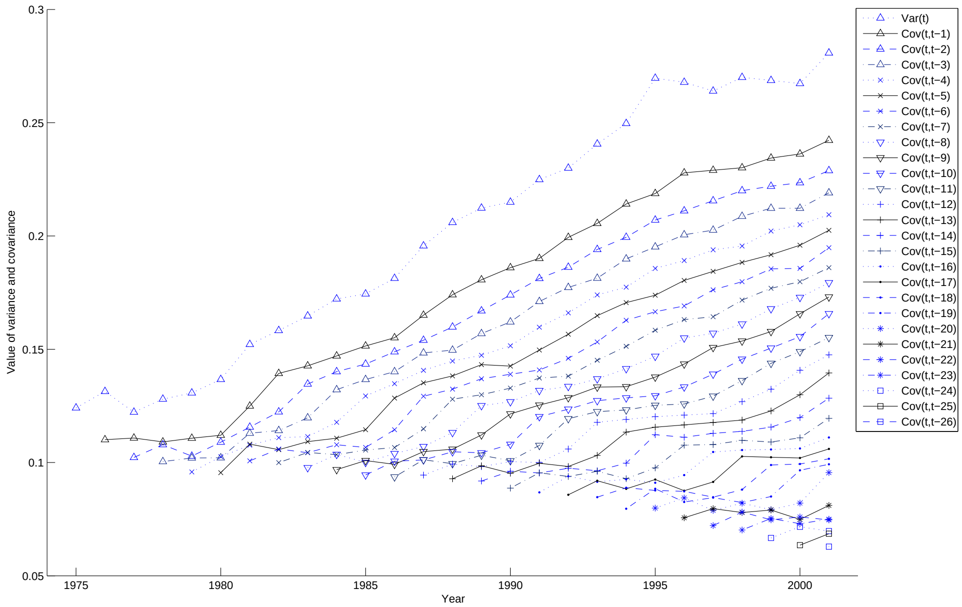

Figure 2: Variance and covariance of log-wage.

3 Methodological Setup

3.1 Calculate Variance-Covariance Matrix of Log-wages

I closely follow the methods in Kalwij and Alessie (2007) to compute the sample variance-covariance matrix of individual wages. In year t, the log-wage of the individual i, who was born in year b, is denoted by $y_{i,b,t}$. For $i \in \{1,\ldots,n\}$, $b \in \{1916,\ldots,1980\}$ and $t \in \{1975,\ldots,2001\}$, $y_{i,b,t}$ can be decomposed into the following two parts:

$$y_{i,b,t} = w_{b,t} + u_{i,b,t}, \tag{1}$$

where $w_{b,t}$ is the population mean of log-wage in year t of the individuals born in year b. I use $\hat{w}_{b,t}$ to denote the corresponding sample average. Let t_b denote the number of observed years for the individuals born in year b. I assume

$$\boldsymbol{u}_{i,b} \sim \text{IID}\left(\boldsymbol{0}, \boldsymbol{\Sigma}_b\right),$$

where $\boldsymbol{u}_{i,b} = (u_{i,b,1}, u_{i,b,2}, \ldots, u_{i,b,t_b})'$ and $\boldsymbol{\Sigma}_b$ is a $t_b \times t_b$ matrix.

Before I explain how I have estimated the elements of $\boldsymbol{\Sigma}_b$, let $\hat{\boldsymbol{u}}_{b,t}$ be defined as follows:

$$\hat{\boldsymbol{u}}_{b,t} = (\hat{u}_{1,b,t}, \hat{u}_{2,b,t}, \ldots, \hat{u}_{i,b,t}, \ldots, \hat{u}_{n_b,b,t})',$$

where n_b is the number of observations for distinct individuals born in year b and

$$\hat{u}_{i,b,t} = \begin{cases} y_{i,b,t} - \hat{w}_{b,t} & \text{if individual } i \text{ is observed in year } t \\ 0 & \text{otherwise.} \end{cases}$$

Similarly, the indicator vector $\boldsymbol{d}_{b,t}$ is defined as

$$\boldsymbol{d}_{b,t} = (d_{1,b,t}, d_{2,b,t}, \ldots, d_{i,b,t}, \ldots, d_{n_b,b,t})',$$

where

$$d_{i,b,t} = \begin{cases} 1 & \text{if individual } i \text{ is observed in year } t \\ 0 & \text{otherwise.} \end{cases}$$

The (k,l)-th element of $\boldsymbol{\Sigma}_b$ is estimated by $\boldsymbol{M}_b(k,l)$:

$$\boldsymbol{M}_b(k,l) = \frac{\hat{\boldsymbol{u}}'_{b,k}\hat{\boldsymbol{u}}_{b,l}}{\boldsymbol{d}'_{b,k}\boldsymbol{d}_{b,l}},$$

with $k,l = \{1,\ldots,t_b\}$. Since we assume that we have a random sample at disposal and the panel attrition is random, by Law of Large Numbers, we can claim that $\boldsymbol{M}_b(k,l)$ is a consistent estimate for $\boldsymbol{\Sigma}_b(k,l)$.

The vector $\boldsymbol{m}_b$ contains the $t_b(t_b+1)/2$ distinct elements of the symmetric matrix $\boldsymbol{M}_b$, and $\boldsymbol{m}_b = \text{VECH}(\boldsymbol{M}_b)$, where VECH is the operation that stacks the distinct elements of a matrix. Furthermore, I define $\boldsymbol{\sigma}_b = \text{VECH}(\boldsymbol{\Sigma}_b)$. Chamberlain (1984) shows that $\boldsymbol{m}_b$ has the following distribution asymptotically:

$$\boldsymbol{m}_b \xrightarrow{d} \mathcal{N}\left(\boldsymbol{\sigma}_b, \boldsymbol{V}_b\right),$$

and the element of the covariance matrix $\boldsymbol{V}_b$ is estimated as follows:

$$\widehat{\boldsymbol{V}}_b\left(\boldsymbol{M}_b(k,l), \boldsymbol{M}_b(p,q)\right) = \frac{\sum_{i=1}^{n_b} d_{i,b,k}d_{i,b,l}d_{i,b,p}d_{i,b,q}}{(\boldsymbol{d}'_{b,k}\boldsymbol{d}_{b,l})(\boldsymbol{d}'_{b,p}\boldsymbol{d}_{b,q})}\left(\boldsymbol{M}_b(k,l,p,q) - \boldsymbol{M}_b(k,l)\boldsymbol{M}_b(p,q)\right),$$

with

$$\boldsymbol{M}_b(k,l,p,q) = \frac{\sum_{i=1}^{n_b} \hat{u}_{i,b,k}\hat{u}_{i,b,l}\hat{u}_{i,b,p}\hat{u}_{i,b,q}}{\sum_{i=1}^{n_b} d_{i,b,k}d_{i,b,l}d_{i,b,p}d_{i,b,q}}$$

where $k, l, p, q = \{1, \cdots, t_b\}$ and $k \leq l, p \leq q$.

Let $\boldsymbol{m}$ denote the vertical concatenation of all the $\boldsymbol{m}_b$ vectors:

$$\boldsymbol{m} = (\mathrm{VECH}(\boldsymbol{M}_{1916})', \ldots, \mathrm{VECH}(\boldsymbol{M}_{1980})')',$$

and define $\boldsymbol{V}$ to be a block diagonal matrix with diagonal $\boldsymbol{V}_{1916}, \ldots, \boldsymbol{V}_{1980}$. In my case, $\boldsymbol{m}$ is a 11466×1 vector and $\boldsymbol{V}$ is a 11466×11466 matrix.

3.2 Theoretical Variance-Covariance Component Model

Having discussed the ways to calculate the sample variance-covariance matrix of log-wage, in this section I introduce the theoretical structure of the individual earnings process. I follow the steps of Guvenen (2009). To begin with, I use the following structure to model the deviation term in Equation (1):

$$u_{i,b,t} = f(\alpha_i, \beta_i, \boldsymbol{X}_{i,b,t}) + v_{i,t} + \phi_t \epsilon_{i,t}. \tag{2}$$

In the above equation, function f models the permanent component (life-cycle profile) of the individual earnings process and the rest part captures the transitory shocks that affect individuals or labor market. $\boldsymbol{X}_{i,b,t}$ is the information set, which may contain the information on individual's age, education level, job category and etc.

The key difference between the HIP and the RIP process concerns with the specification of function f. In the HIP process, it is assumed that each person has an individual-specific earnings profile. If the growth rate of wages depends on individual's ability, education level and occupation, this dependence will be captured by the individual-specific parameters in f. The HIP process in Guvenen (2009) assumes the following first order linear specification in working experience to model the individual-specific life-cycle profile:

$$f(\alpha_i, \beta_i, \boldsymbol{X}_{i,b,t}) = \alpha_i + \beta_i g, \tag{3}$$

where g is individual i's years of working experience in year t, α_i and β_i are jointly iid distributed across individuals with zero mean, variance σ_α^2 and σ_β^2 respectively, and covariance $\sigma_{\alpha\beta}$. Although it is possible to use second or higher order linear specification, Baker (1997) discusses that it is unnecessary to extend the first order linear specification to higher order, since the parameter estimates and the fit of the model do not improve dramatically.

On the contrary, in the RIP process all individuals face a similar life-cycle profile, which is simply modeled by an individual fixed-effect α_i:

$$f(\alpha_i, \beta_i, \boldsymbol{X}_{i,b,t}) = \alpha_i. \tag{4}$$

By the above specifications, it is obvious that the RIP process is nested in the HIP process. In addition, in the RIP process it is also possible to model the permanent component using a more complicated structure. For example, to allow the term α_i to change with age, Dickens (2000) and Kalwij and Alessie (2007) use a random walk in age with age-varying innovation variance to model this term. This specification is not necessary for the present aim of this study.

Similar as Guvenen (2009), I use an AR(1) process and a purely transitory shock to model the dynamic component of wages,

$$v_{i,t} = \rho v_{i,t-1} + \pi_t \eta_{i,t}, \quad \text{with} \quad \eta_{i,t} \sim \text{iid}(0, \sigma_\eta^2). \tag{5}$$

To permit the time effects and possible time-variations in the innovations of the AR(1) part, I multiply the innovation term of each year by the corresponding π_t's. Analogously, the component $\epsilon_{i,t}$ is also iid distributed with mean zero and variance σ_ϵ^2, and the possible time effects in this term are formalized by ϕ_t's. According to Bound and Krueger (1991), if the measurement errors are independent over time, they will be included in $\phi_t \epsilon_{i,t}$. Otherwise, the measurement errors will be captured by the AR(1) part if they are auto-correlated.

Given the specifications in Equation (2), (3) and (5), it is straightforward to obtain the theoretical variance-covariance structure of the HIP process:

$$\text{var}(u_{i,b,t}) = \left[\sigma_\alpha^2 + (g+g)\sigma_{\alpha\beta} + g^2\sigma_\beta^2\right] + \text{var}(v_{i,t}) + \phi_t^2\sigma_\epsilon^2 \tag{6}$$

$$\text{cov}(u_{i,b,t}, u_{i,b,s}) = \left[\sigma_\alpha^2 + (g+h)\sigma_{\alpha\beta} + gh\sigma_\beta^2\right] + \rho^{h-g}\,\text{var}(v_{i,t}), \tag{7}$$

where g and h are years of working experience in year t and s ($s > t$) respectively. In this study, the working experience g and h are approximated by the difference between his/her age in year t and age twenty-one: $g = t - b - 20$, $h = s - b - 20$. Although it is ideal to use a measure of actual working experience, this information is not available in the NES. For the theoretical variance-covariance structure of the RIP process, $\sigma_{\alpha\beta}$ and σ_β^2 in Equation (6) and (7) are restricted to be zero.

If the individuals born in year b have one year working experience in their first observed year, the variance of the AR(1) part in this year is calculated by:

$$\text{var}(v_{i,t}) = \pi_t^2\sigma_\eta^2. \tag{8}$$

And if the individuals born in year b have more than one year working experience in the first observation year (year 1975), the variance of the AR(1) part in year 1975 is calculated by

$$\operatorname{var}(v_{i,t}) = \pi_t^2 \sigma_\eta^2 \sum_{j=0}^{g-1} \rho^{2j}, \quad \text{if} \quad g > 1, t = 1975. \tag{9}$$

The variance of the AR(1) part is calculated recursively:

$$\operatorname{var}(v_{i,t}) = \rho^2 \operatorname{var}(v_{i,t-1}) + \pi_t^2 \sigma_\eta^2. \tag{10}$$

Equation (9) assumes that the coefficients of the innovation term of the AR(1) part are constant before the survey started in year 1975. For the individuals who have g ($g > 1$) year's working experience in the first observation year (year 1975), the variance of the AR(1) part in year 1975 is accumulated over the last $g-1$ years. It is obvious that ϕ_t only has an instantaneous effect on the cross-sectional variance at time t, whereas π_t has lasting effects on the subsequent variance-covariance. To obtain identification, I normalize ϕ_t and π_t to be one for the first observation year ($\phi_{1975} = 1$ and $\pi_{1975} = 1$) and impose the restriction $\pi_{2000} = \pi_{2001}$.

Kalwij and Alessie (2007) and Doris et al. (2010) include cohort effects non-parametrically in both the permanent and transitory components. To incorporate cohort factor loadings into the theoretical structure of Guvenen (2009), I extend Equation (6) and (7) in the following way:

$$\operatorname{var}(u_{i,b,t}) = r_b^2 \left[\sigma_\alpha^2 + (g+g)\sigma_{\alpha\beta} + g^2\sigma_\beta^2\right] + s_b^2 \left[\operatorname{var}(v_{i,t}) + \phi_t^2 \sigma_\epsilon^2\right] \tag{11}$$

$$\operatorname{cov}(u_{i,b,t}, u_{i,b,s}) = r_b^2 \left[\sigma_\alpha^2 + (g+h)\sigma_{\alpha\beta} + gh\sigma_\beta^2\right] + s_b^2 \left[\rho^{h-g} \operatorname{var}(v_{i,t})\right]. \tag{12}$$

Theoretically, the above model is identified if r_b and s_b are normalized to be one for the 1916 and 1917 birth cohorts and r_b's and s_b's are equalised for the 1979 and 1980 birth cohorts. In practice, the cohort effects are poorly estimated and the corresponding standard errors are extremely large. For this reason I normalise r_b and s_b to be one for the birth cohorts 1916 to 1924, and equalise r_b's and s_b's for the birth cohorts 1978 to 1980.

Given the above specifications, the parameters to be estimated are $\boldsymbol{b} = [\sigma_\alpha^2, \sigma_{\alpha\beta}, \sigma_\beta^2, \sigma_\eta^2, \sigma_\epsilon^2, \pi_{1975}, \ldots, \pi_{2001}, \phi_{1975}, \ldots, \phi_{2001}, r_{1916}, \ldots, r_{1980}, s_{1916}, \ldots, s_{1980}]$.

3.3 Minimum Distance Estimation

The parameters of the theoretical earnings process are estimated by applying minimum distance (MD) estimation, which provides us a computationally convenient method for fitting a mathematical model to data without the need for distributional assumptions. It is first proposed by Chamberlain (1984) and since then has been widely applied to a range of problems.

I fit the vector $\boldsymbol{m}$ to the vector function $\boldsymbol{f}$ specified by Equation (6) and (7). The minimum distance estimator $\hat{\boldsymbol{b}}_{\text{MD}}$ minimizes the following quadratic objective function with respect to $\boldsymbol{b}$:

$$Q(\boldsymbol{b}) = [\boldsymbol{m} - \boldsymbol{f}(\boldsymbol{b})]' \boldsymbol{A} [\boldsymbol{m} - \boldsymbol{f}(\boldsymbol{b})],$$

where $\boldsymbol{A}$ is the weighting matrix. And

$$\hat{\boldsymbol{b}}_{\text{MD}} = \arg\min [\boldsymbol{m} - \boldsymbol{f}(\boldsymbol{b})]' \boldsymbol{A} [\boldsymbol{m} - \boldsymbol{f}(\boldsymbol{b})].$$

Chamberlain (1984) discusses that the optimal choice for the weighting matrix $\boldsymbol{A}$ is $\boldsymbol{V}^{-1}$. However, Altonji and Segal (1996) provide a discussion of the bias issue that arises in minimum distance estimation. They examine the small sample property of minimum distance estimation based on a Monte Carlo simulation study, and show that the Optimal Weighting Minimum Distance (OWMD) estimates lead to serious sample bias due to correlation between the sampling errors in $\boldsymbol{m}$ and $\boldsymbol{V}$. Furthermore they suggest using Equally Weighted Minimum Distance (EWMD), where an identity matrix $\boldsymbol{I}$ is used instead of $\boldsymbol{V}^{-1}$. In my estimation I follow this suggestion, as the majority of the literature do.

Under suitable regularity conditions, the MD estimator $\hat{\boldsymbol{b}}_{\text{MD}}$ is consistent, asymptotically normal distributed with the following covariance matrix

$$\boldsymbol{V}(\hat{\boldsymbol{b}}_{\text{MD}}) = (\boldsymbol{G}'\boldsymbol{A}\boldsymbol{G})^{-1}\boldsymbol{G}'\boldsymbol{A}\boldsymbol{V}\boldsymbol{A}\boldsymbol{G}(\boldsymbol{G}'\boldsymbol{A}\boldsymbol{G})^{-1}, \tag{13}$$

where $\boldsymbol{G}$ is the Jacobian matrix of the structural function $\boldsymbol{f}$ evaluated at $\hat{\boldsymbol{b}}_{\text{MD}}$,

$$\boldsymbol{G} = \frac{\partial \boldsymbol{f}(\boldsymbol{b})}{\partial \boldsymbol{b}} |\hat{\boldsymbol{b}}_{\text{MD}}. \tag{14}$$

$\boldsymbol{G}$ is a $n \times k$ matrix, where n is the length of $\boldsymbol{m}$ and k is the length of $\boldsymbol{b}$.

The minimum distance estimation also provides us with a χ^2 test to evaluate the goodness of fit between the estimated $\boldsymbol{f}(\hat{\boldsymbol{b}}_{\text{MD}})$ and $\boldsymbol{m}$:

$$\left[\boldsymbol{m} - \boldsymbol{f}(\hat{\boldsymbol{b}}_{\text{MD}})\right] \boldsymbol{R}^{-} \left[\boldsymbol{m} - \boldsymbol{f}(\hat{\boldsymbol{b}}_{\text{MD}})\right]' \sim \chi^2(q), \tag{15}$$

with $\boldsymbol{R} = \boldsymbol{W}\boldsymbol{V}\boldsymbol{W}'$, $\boldsymbol{W} = \boldsymbol{I} - \boldsymbol{G}\left(\boldsymbol{G}'\boldsymbol{A}\boldsymbol{G}\right)^{-1}\boldsymbol{G}'\boldsymbol{A}$. $\boldsymbol{R}^{-}$ is a generalised inverse of $\boldsymbol{R}$. The degrees of freedom q equal $n - k$. In this study, n equals 11466, k equals 59 or 189, depending on whether I estimate cohort effects or not.

3.4 The Detailed Information of Estimation Procedure

In this part I discuss the detailed information of my estimation procedure, as well as the problems and issues I met. For readers who are only interested in the main estimation results, they may skip this part.

The optimization procedure is performed on MATLAB 7.10.0, installed on the student PC offered by the University of Groningen[2]. Two optimization functions `fmincon` and `lsqnonlin` in the MATLAB Optimization Toolbox(TM) are used to search for the minimum distance estimator[3]. For the optimization function `fmincon`, the converging time to obtain the optimal solution is approximately 50 minutes (for the specifications in Equation (6) and (7)) and 80 minutes (for the specifications in Equation (11) and (12)); for the optimization function `lsqnonlin`, the converging time to obtain the optimal solution is approximately 30 minutes and 60 minutes respectively.

The Jacobian Matrix $\boldsymbol{G}$ is approximated by using finite differences. In case of Equally Weighted Minimum Distance estimation (when $\boldsymbol{A} = \boldsymbol{I}$), the covariance matrix $\boldsymbol{V}(\hat{\boldsymbol{b}}_{\text{MD}})$ defined in Equation (13) cannot be computed directly due to the memory limitation of the PC. However, I can partition each matrix in Equation (13) by cohorts and perform the calculation on the basis of each partitioned matrix. See Appendix A for more information. Netherless, due out of memory problem, the Optimal Weighting Minimum Distance (OWMD) estimates, matrix $\boldsymbol{W}$, matrix $\boldsymbol{R}$ and the goodness of fit test in Equation (15) cannot be calculated by using the PC that is available for me [4].

To solve the out of memory problem and decrease the converging time of optimization, one may perform the calculation on the Windows 64-bits system. A more advanced solution is suggested by using the MATLAB Parallel Computing Toolbox(TM), which allows us to solve computationally and data-intensive problems using multicore processors, GPUs, and computer clusters. Furthermore, the MATLAB Parallel Computing Toolbox(TM) also provides the possibility to launch the "cloud" calculations on the MATLAB Distributed Computing Server running on Amazon EC2. Due to a number of reasons, I did not to try either of these two solutions.

---

[2]Operation system: Windows XP, 32-bits. CPU: Intel(R) Core(TM) 2 DUO CPU, E8400@3.00GHZ. L1-cache: 32K(8-way set associative, 64-byte line size). Memory: 1992MB/sec sustained transfer rate.

[3]`lsqnonlin` is special designed to solves nonlinear least-squares curve fitting problems of the form: $\min ||\boldsymbol{f}(\boldsymbol{x})||_2^2 = \min_{\boldsymbol{x}}(f_1(\boldsymbol{x})^2 + f_2(\boldsymbol{x})^2 + \cdots + f_n(\boldsymbol{x})^2)$. This is exactly the form of my objective function in case of using Equally Weighted Minimum Distance (EWMD) estimation.

[4]I even met the out of memory problem when I generate the 11466×11466 matrix $\boldsymbol{I}$, which is used to calculate the matrix $\boldsymbol{W}$. Approximately matrix $\boldsymbol{I}$ occupies 1 GB memory.

4 Empirical Results

4.1 Evidence for the RIP Process

Panel A of Table 2 reports the main estimation results of the HIP process that allows for individual-specific experience earnings profile. In row 1, I follow the way of Guvenen (2009) and estimate σ^2_α, σ^2_β and $corr_{\alpha,\beta}$ (with the restriction $-1 \leq corr_{\alpha,\beta} \leq 1$) in Equation (6) and (7). By using $corr_{\alpha,\beta} = \sqrt{\sigma_{\alpha,\beta}}/\sigma_\alpha\sigma_\beta$, I can calculate $\hat{\sigma}_{\alpha,\beta}$. To rule out other possibilities, in row 2 I also estimate σ^2_α, σ^2_β and $\sigma_{\alpha,\beta}$ directly and calculate $\hat{corr}_{\alpha,\beta}$. All in all, the heterogeneity parameters (σ^2_β, $\sigma_{\alpha,\beta}$ and $corr_{\alpha,\beta}$) are extremely close to zero and not significant at conventional confidence levels. The estimates ρ's are around 0.980. In comparison with $\rho = 0.82$ in the HIP process of Guvenen (2009), the estimated parameter ρ is strongly persistent and does not support for the HIP process either. Therefore, in contrast with Guvenen (2009), I cannot find evidence for the validity of the HIP process in the life-cycle profile of wages in the UK.

To estimate the RIP process that ignores heterogeneity in life-cycle profile, I impose the restrictions $\sigma^2_\beta = 0$ and $\sigma^2_{\alpha,\beta} = 0$ on Equation (6) and (7). The estimates of the main parameters of interest are reported in row 3 of Table 2. First notice that the estimated persistence parameter ρ is 0.980. This value is very close to the estimate of the RIP process in Guvenen (2009)($\rho = 0.988$) and the estimates in other studies that use the RIP process, such as Dickens (2000) ($\rho = 0.973$) and Kalwij and Alessie (2007) ($\rho = 0.980$). In terms of other parameters, the variance of non-heterogeneity term σ^2_α, the innovation variance of the AR(1) part σ^2_η and the variance of the purely transitory shocks σ^2_ϵ are significant at conventional confidence levels, and they do not differ dramatically from those in row 1 and 2 of the HIP process. In addition, using Equation (11) and (12), row 4 repeats the estimation procedure of row 3 while controlling for the birth cohort effects. Including the cohort effects yields an improvement in the goodness of fit, and the estimated coefficients of the cohort effects are significant. Other estimation results can be found in Table 3 and Table 4 in Appendix B.

My sample selection criteria is another informal support which makes my estimation results in favor of the RIP process. As discussed in Section 2.2, Guvenen (2009) selects the individuals that are included in the sample for at least twenty years. He also draws a new sample for his estimation by imposing a release requirement that individuals have to stay in the sample for at least four years (not necessarily consecutive), so his new sample contains more heterogeneous individuals. As can be anticipated, the estimated results of the new sample show stronger evidence of the heterogeneity parameters (σ^2_β and $corr_{\alpha,\beta}$). In this paper, I do not put a similar restriction on the sample selection procedure. One can infer that the evidence of the HIP process would be even less if I put a similar restriction.

	HIP/RIP	σ^2_α	$\sigma^2_{\alpha,\beta}$	$corr_{\alpha,\beta}$	σ^2_β	σ^2_η	σ^2_ϵ	ρ	criteria fun.
Panel A									
(1)	HIP	0.047	1.482×10^{-4}	0.999	$4.640 \times 10^{-7*}$	0.005	0.008	0.977	2.723
		(9.815×10^{-4})	/	(5.260)	(4.814×10^{-6})	(2.110×10^{-4})	(1.408×10^{-3})	(2.221×10^{-3})	
		48.232	/	*0.190*	*0.096*	*22.767*	*5.388*	*440.084*	
(2)	HIP	0.046	2.959×10^{-4}	1387.316*	9.995×10^{-13}	0.005	0.009	0.975	2.709
		(9.918×10^{-4})	(6.1974×10^{-5})	/	(4.535×10^{-6})	(2.167×10^{-4})	(1.414×10^{-3})	(2.235×10^{-3})	
		45.880	*4.774*	/	2.204×10^{-7}	*21.186*	*6.106*	*436.317*	
Panel B									
(3)	RIP	0.049	0	0	0	0.005	0.007	0.980	2.739
		(7.007×10^{-4})	/	/	/	(2.092×10^{-4})	(1.684×10^{-3})	(1.112×10^{-3})	
		70.006	/	/	/	*24.061*	*3.915*	*879.037*	
(4)	RIP	0,035	0	0	0	0,004	0.008	0.990	1.012
		(2.105×10^{-3})	/	/	/	(3.118×10^{-4})	(8.635×10^{-4})	(9.413×10^{-4})	
		16.788	/	/	/	*11.796*	*8.948*	*1050.270*	

Note: 1. In row (1) I estimate σ^2_β and $corr_{\alpha,\beta}$; In row (2) I estimate σ^2_β and $\sigma_{\alpha,\beta}$; In row (3) I restrict $\sigma^2_\beta = 0$ and $\sigma^2_{\alpha,\beta} = 0$ (the RIP process); In row (4) I repeat the same estimation procedure of row (3) by controlling the cohort effects.
2. Standard errors are in parentheses; slanted number is t-value.
3. Sign * means I estimate this variable by calculation through: $corr_{\alpha,\beta} = \sqrt{\sigma_{\alpha,\beta}}/\sigma_\alpha\sigma_\beta$.
4. The detailed estimation results of the time and cohort effects are reported in Appendix B.

Table 2: Results of the HIP and RIP processes

To obtain a better insight of the time effects, Figure 3 plots the estimated variance of the innovation term in the AR(1) part by using the estimates in row 4.[5] Except two peaks in year 1976 and 1981, it has a slightly increasing trend over time. Recall that the π_t in 1975 is normalized to be one, the large innovation variance in year 1976 is mainly attributed to the Secondary Banking Crisis of 1973-1975 in the United Kingdom. The estimated variance of the term $\phi_t \epsilon_{i,t}$ is plotted in Figure 4. The variance increases substantially since 1977 and rises sharply from year 1992 to 1995. After peaking in year 1995, it decreases from then on [6].

Assuming $r_b = 1$ and $s_b = 1$ for all b, Figure 5 depicts the estimated variance of the transitory component $(v_{i,t} + \phi_t \epsilon_{i,t})$. The variance of the transitory component increases since year 1975 and levels off after year 1995, which is consistent with the descriptive finding that there is a massive increase in wage inequality before year 1995. This finding is in contrast with Dickens (2000) but line with Kalwij and Alessie (2007), who attribute this massive increase mainly to a strong rise in transitory wage inequality.

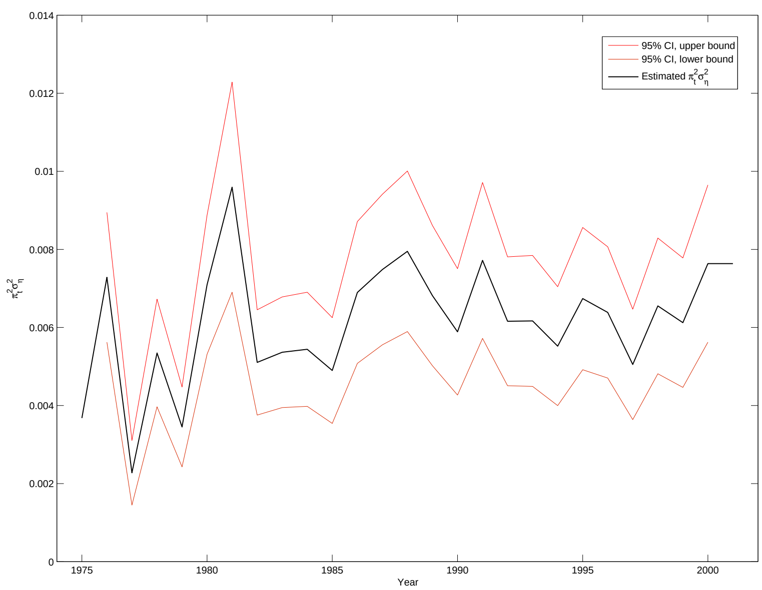

Figure 3: The variance of the innovation term in the AR(1) part (Estimated $\pi_t^2 \sigma_\eta^2$ in Equation (10)).

[5]The estimates in row 3 imply a similar time trend. The estimates in row 3 also procedure similar time trends in Figure 4 and 5.

[6]The null hypotheses for no time effects are $H_0 : \pi_{1976}^2 = \cdots = \pi_{2000}^2 = 1$ and $H_0 : \phi_{1976}^2 = \cdots = \phi_{2001}^2 = 1$. The corresponding $\chi^2(25)$ and $\chi^2(26)$ test statistics equal 1370 (p-value=0.000) and 13819 (p-value=0.000) respectively.

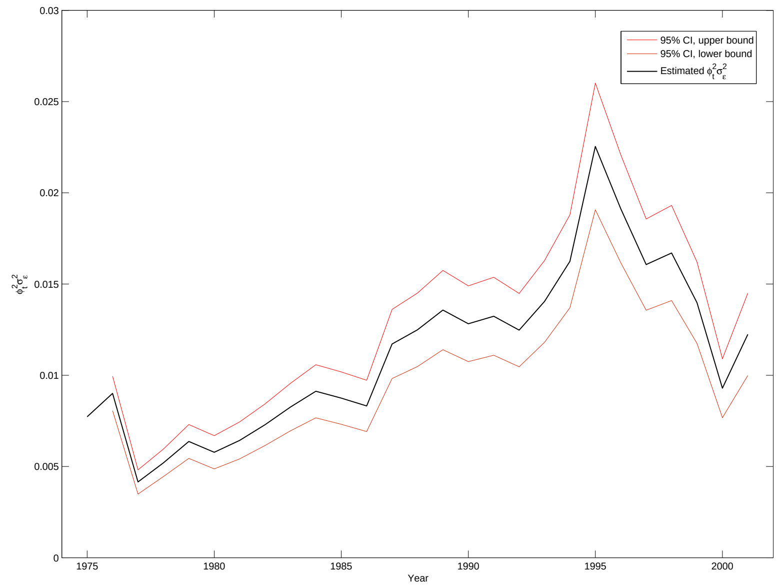

Figure 4: The variance of purely transitory shocks (Estimated $\phi_t^2\sigma_\epsilon^2$ in Equation (11)).

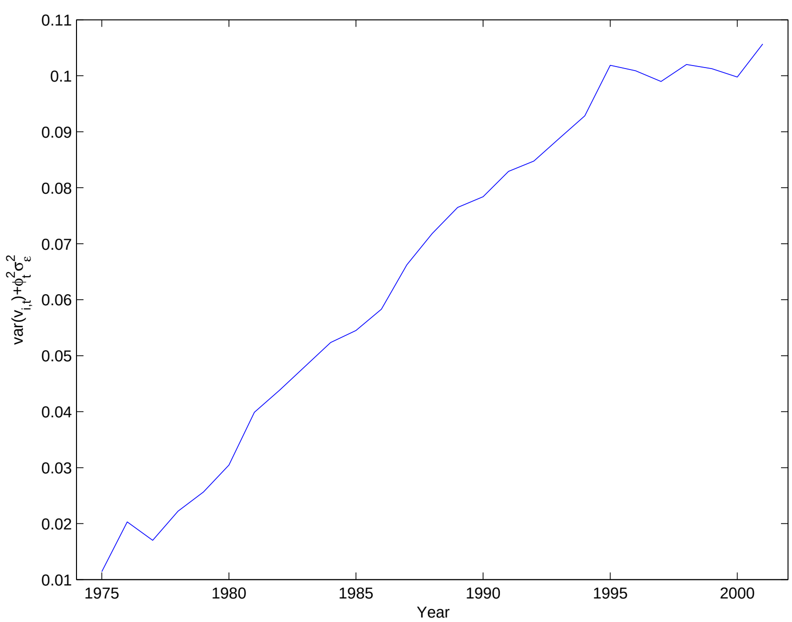

Figure 5: The variance of the transitory component (predicted $\mathrm{var}(v_{i,t}) + \phi_t^2\sigma_\epsilon^2$).

4.2 AR(1) Part and its Accumulated Effects

The theoretical structural specifications in Section 3.2 imply that if there exists an autocorrelation in measurement error and the measurement error has a lower persistence than income shocks, the estimated ρ will be smaller than the true persistence parameter of income shocks. However, such an under-estimation does not seem to exist, since the estimates are every closed to one in both RIP and HIP specifications. In Figure 6, I plot the remaining effect of an AR(1) shock after thirty years for different values of the persistence parameter ρ. My estimate $\rho = 0.980$ implies that the impact of an income shock upon earnings process is strongly persistent. The effect of an income shock remains more than fifty percent of its initial value even after thirty years. In case of modest persistent income shocks, for example when $\rho = 0.82$ (Guvenen, 2009), this effect is reduced to fifty percent of its initial value after four years and almost fades out after twenty years. For convenience of comparison, I also plot the response for $\rho = 1$ and $\rho = 0.64$ (Haider, 2001). Individuals in the UK and the US are likely to face two types of income shocks and therefore they may tend to make different consumption and saving choices.

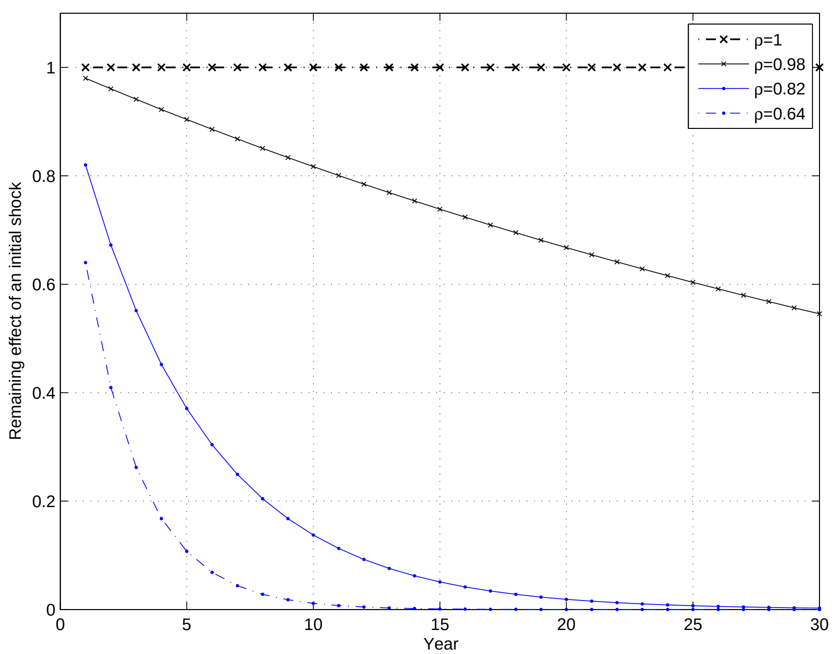

Figure 6: Remaining effect of an initial AR(1) shock

By substituting Equation (10) into (6) and setting $\pi_t = 1$ and $\phi_t = 1$ for all t, Guvenen

(2009) shows that Equation (6) can be expressed as:

$$\mathrm{var}(u_{i,b,t}) = \left(\sigma_\alpha^2 + \sigma_\epsilon^2\right) + \left(\frac{(1-\rho^{2g+1})\sigma_\eta^2}{1-\rho^2}\right) + \left(2g\sigma_{\alpha\beta} + g^2\sigma_\beta^2\right).$$

Since the heterogeneity terms are not significant, the above equation becomes:

$$\mathrm{var}(u_{i,b,t}) = \left(\sigma_\alpha^2 + \sigma_\epsilon^2\right) + \left(\frac{(1-\rho^{2g+1})\sigma_\eta^2}{1-\rho^2}\right).$$

This equation shows that the cross-sectional variance of wage inequality can be decomposed into two parts. The first part in the parentheses contains the terms that are independent of age, and the second part in the parentheses is the accumulated effect of the AR(1) part. Figure 7 plots these two parts by using the estimation results of the RIP process in row 3 of Table 2. Since the terms in the first parentheses are independent of age, altogether their effects on cross-sectional variance remain at a constant level. It is also obvious that the accumulated effect of the AR(1) part is a monotonic increasing concave function of age.

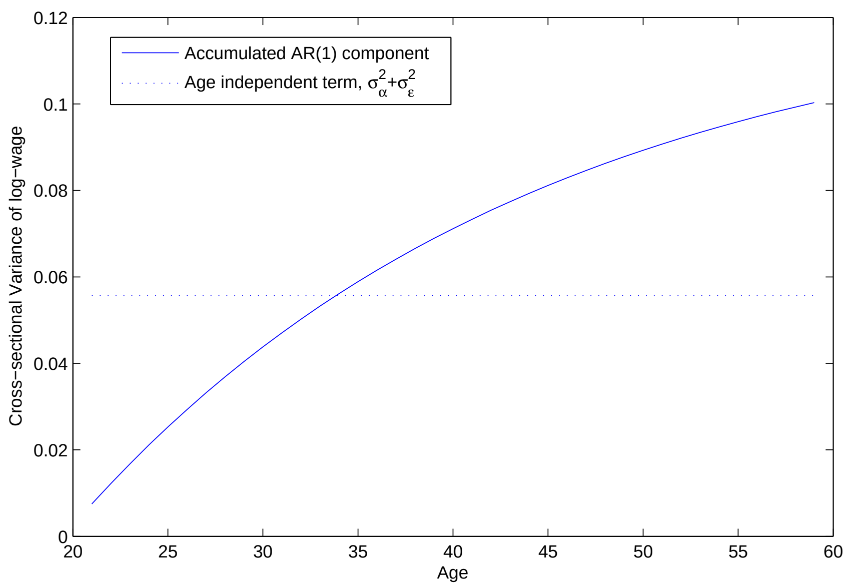

Figure 7: Accumulated effect of the AR(1) part.

4.3 Age-variance Profile

Excluding the insignificant heterogeneity parameters, I rewrite Equation (6) as:

$$\mathrm{var}(u_{i,b,t}) = \sigma_\alpha^2 + \mathrm{var}(v_{i,t}) + \phi_t^2 \sigma_\epsilon^2. \tag{16}$$

Recall that the accumulated effect of the AR(1) part is an increasing concave function of age (under the condition $\rho < 1$). By the above specification, if the within-cohort cross-sectional variance increases with age in a concave way, it will be captured by the auto-correlated AR(1) part. Otherwise, the changes of the cross-sectional variance will be captured by the non-heterogeneity life-cycle profile σ_α^2 and the purely transitory shocks $\phi_t^2\sigma_\epsilon^2$.

Figure 1 in Section 2.4 shows the concavity in the age-variance profile, given some informal support to the concave shape between the cross-sectional variance and age. Ignoring all the time effects and assuming a single cohort for all individuals, I plot the age-variance profile in Figure 8 on the basis of the raw data. The dark line in Figure 8 clearly shows that the cross-sectional variance increases with age in a concave fashion. However, it is also important to take into account the time effects and the cohort effects. Hall (1971) points out the difficulty in identifying the age, time and cohort effects at the same time, since any of these three types of effects can be written as a linear combination of the other two. Feasible approaches to control these effects are discussed by Deaton and Paxson (1994) and Guvenen (2009), who use a non-parametric specification and mainly base on raw data.[7] The light line in Figure 8 plots the age-variance profile after controlling the cohort effects by using the approach of Guvenen (2009). As can be seen from the figure, the light line shows a slight concavity in the age-variance profile and predicts larger variance after age forty. Nevertheless, my results appear to be very strange when I use this approach to control the time effects, so I will not present or discuss them here. [8]

An alternative approach to control these effects is suggested by Kalwij and Alessie (2007). Based on their theoretical model and results, they estimate the age-variance, time-variance and cohort-variance profile in a more parametric way. Similar to my finding but in contrast with in Guvenen (2009), they also find that in the UK the cross-sectional variance of log-wage increases with individuals' life-cycle in a concave fashion. Since my results do not show any evidence for the HIP process but support for the RIP process, the model in Kalwij and Alessie (2007) which assumes the RIP process, is basically an extension of the specification in this study. The findings in Kalwij and Alessie (2007) and in this paper

[7] The detailed steps to control cohort effects of Guvenen (2009) are the following: first construct overlapping age interval $[g-2, g+2]$ for each middle point g (g ranges from 23 to 57 in my case). Then group all individuals based on the combinations of each overlapping age interval and each cohort and computer the variance for each combination. Next regress these variance on a full set of age and cohort dummies, and the coefficients of the age dummies are the estimates we want to obtain. To control the time effects, regress these variances on age and time dummies.

[8] There are some negative values in my estimates results.

both conclude that the age-variance profile is concave in the UK.

The distinction in the age-variance profile between the UK and the US is predominantly caused by whether the life-cycle profile is individual-specific or similar for all individuals. In case of the HIP process, the heterogeneity terms $\left[(g+h)\sigma_{\alpha\beta} + g^2\sigma_\beta^2\right]$ in Equation (6) is a convex function of age. The cross-sectional variance increases with age in a convex fashion due to the presence of these terms (Guvenen, 2009). Netherless, if the heterogeneity terms do not exist, the variance is mainly captured by the AR(1) part, resulting in the concavity in the age-variance profile.

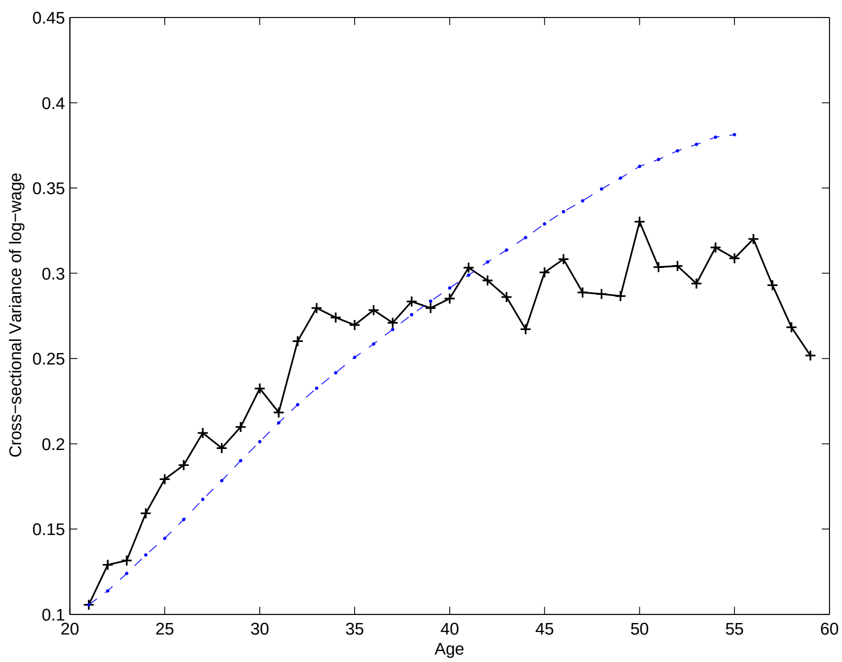

Figure 8: Age-variance profile of log-wage.

4.4 Age-covariance Profile

After removing the insignificant heterogeneity terms, I rewrite the theoretical covariance specification in Equation (7) as:

$$\text{cov}(u_{i,g}, u_{i,h}) = \sigma_\alpha^2 + \rho^{(h-g)}\,\text{var}(v_{i,g}). \tag{17}$$

I drop subscript b since I assume all individuals are from the same cohort. t and s is replaced by g and h since they are perfectly correlated in this case. Replacing h by $h+1$ and taking difference on both sides of Equation (17), the following equation can be obtained:

$$\operatorname{cov}(u_{i,g}, u_{i,h+1}) - \operatorname{cov}(u_{i,g}, u_{i,h}) = \rho^{(h-g)}(\rho - 1)\operatorname{var}(v_{i,g}). \quad (18)$$

Equation (17) implies that the covariance between age g and h can be decomposed into two parts. The first part captures the life-cycle effect, which is independent of age in my specification. The second term captures the effect of the auto-correlated AR(1) part, which entirely depends on h through the discounting term ρ^{h-g} for a fixed g. Given a fixed age g, the covariance of log-wage between age g and h decreases towards zero at the same rate for each increasing value of h. Since the estimated ρ in the RIP process is 0.980, the covariance decreases almost linearly in h.

Equation (18) calculates the difference (slope) between two consecutive covariances $\operatorname{cov}(u_{i,g}, u_{i,h+1})$ and $\operatorname{cov}(u_{i,g}, u_{i,h})$. Since ρ remains at 0.980, the difference is determined by the age difference $h-g$, and the initial variance of AR(1) part at age g, $\operatorname{var}(v_{i,g})$. Since the specification in Equation (9) assumes that $\operatorname{var}(v_{i,g})$ increases with g, as individuals get older, the age-covariance profile starts with a higher value and decreases faster. The above theoretical discussions are summarized in Figure 9 by using the empirical results in row 3 of Table 2.

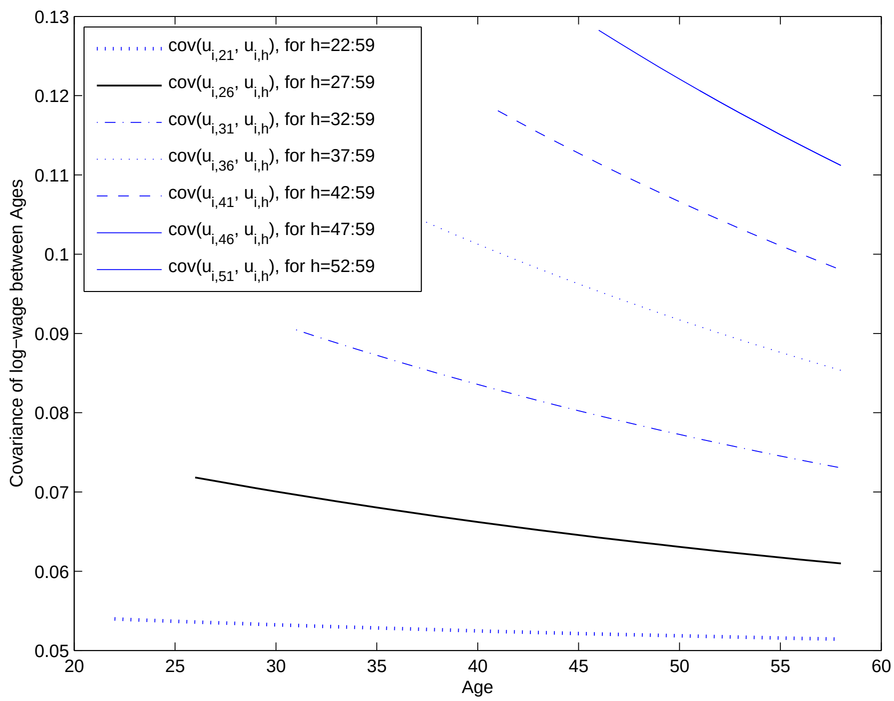

Figure 9: Age-covariance profile of log-wage.

5 Conclusion

The existing studies use either a Restricted Income Profile (RIP) process or Heterogeneous Income Profile (HIP) process to model the individual earnings process. Using the New Earnings Survey panel dataset over the period 1975 to 2001, this paper analyses the variance-covariance structure of individual wage rates in the UK and focuses on examining the evidence for the validity of the RIP process versus the HIP process. I find evidence in favor of the RIP process, providing statistical support to the previous studies that use the RIP process to model the individual earnings process in the UK.

After presenting the general method to calculate the sample variance-covariance matrix of log-wage, I apply minimum distance estimation to fit a theoretical error components model to the sample variance-covariance matrix of log-wage. The empirical results imply that individuals in the UK are subject to large and persistent income shocks and face similar life-cycle profiles (RIP process). However it has been found that individuals in the US are subject to income shocks with modest persistence while facing individual-specific earnings profiles (HIP process). The estimated coefficient of the persistent income shocks implies that the impact of an income shock upon individual earnings process is strongly persistent in the UK. In terms of the RIP and HIP processes, difference also emerges in the age-variance and age-covariance profiles. The evidence of the RIP process leads to the concave shape in age-variance profile, which is in contrast with the convex shape derived by using the PSID dataset of the US.

One possible explanation for the lower persistent income shocks in the US is that the labor market is more flexible and there are more job turnovers in the US. Furthermore, since the PSID dataset used in Guvenen (2009) is not an administrative dataset, there may exist uncertainty that whether the findings from the PSID give accurate results or not. In addition to further research it would be informative to use other administrative datasets to examine whether or not the RIP process holds across other European countries. It also would be interesting to conduct more analyses on the underlying reasons for the evidences of these two processes.

Appendix A

As discussed in Section 3.4, the covariance matrix $\boldsymbol{V}(\hat{\boldsymbol{b}}_{\text{MD}})$ defined in Equation (13) cannot be computed directly due to the memory limitation of the PC. To solve this problem, I partition matrix $\boldsymbol{G}$ by cohort:

$$\boldsymbol{G} = (\boldsymbol{G}_1, \cdots, \boldsymbol{G}_{65})'.$$

In case of using Equally Weighted Minimum Distance estimation, $\boldsymbol{A} = \boldsymbol{I}$, we have

$$\begin{aligned} \boldsymbol{V}(\hat{\boldsymbol{b}}_{\text{MD}}) &= (\boldsymbol{G}'\boldsymbol{A}\boldsymbol{G})^{-1}\boldsymbol{G}'\boldsymbol{A}\boldsymbol{V}\boldsymbol{A}\boldsymbol{G}(\boldsymbol{G}'\boldsymbol{A}\boldsymbol{G})^{-1} \\ &= (\boldsymbol{G}'\boldsymbol{G})^{-1}\boldsymbol{G}'\boldsymbol{V}\boldsymbol{G}(\boldsymbol{G}'\boldsymbol{G})^{-1} = \end{aligned}$$

$$\left[(\boldsymbol{G}_1', \cdots, \boldsymbol{G}_{65}')\begin{pmatrix} \boldsymbol{G}_1 \\ \vdots \\ \boldsymbol{G}_{65} \end{pmatrix}\right]^{-1} (\boldsymbol{G}_1', \cdots, \boldsymbol{G}_{65}')\begin{pmatrix} \boldsymbol{V}_1 & & \boldsymbol{0} \\ & \ddots & \\ \boldsymbol{0} & & \boldsymbol{V}_{65} \end{pmatrix}\begin{pmatrix} \boldsymbol{G}_1 \\ \vdots \\ \boldsymbol{G}_{65} \end{pmatrix}\left[(\boldsymbol{G}_1', \cdots, \boldsymbol{G}_{65}')\begin{pmatrix} \boldsymbol{G}_1 \\ \vdots \\ \boldsymbol{G}_{65} \end{pmatrix}\right]^{-1}$$

$$= \left(\boldsymbol{G}_1'\boldsymbol{G}_1 + \cdots + \boldsymbol{G}_{65}'\boldsymbol{G}_{65}\right)^{-1}\left(\boldsymbol{G}_1'\boldsymbol{V}_1\boldsymbol{G}_1 + \cdots + \boldsymbol{G}_{65}'\boldsymbol{V}_{65}\boldsymbol{G}_{65}\right)\left(\boldsymbol{G}_1'\boldsymbol{G}_1 + \cdots + \boldsymbol{G}_{65}'\boldsymbol{G}_{65}\right)^{-1}.$$

By the above transformation, $\boldsymbol{V}(\hat{\boldsymbol{b}}_{\text{MD}})$ is calculated on the basis of $\boldsymbol{G}_b$ and $\boldsymbol{V}_b$, $b = 1, \cdots, 65$. It is not a problem for the PC to do the multiplication and calculate the inverse for the sizes of $\boldsymbol{G}_b$ and $\boldsymbol{V}_b$.

Appendix B

π_t^2	Estimate	SE	t-value	ϕ_t^2	Estimate	SE	t-value
1975	1.0000			1975	1.0000		
1976	1.9807	0.3048	6.4980	1976	1.1656	0.0744	15.6706
1977	0.6189	0.1367	4.5264	1977	0.5375	0.0522	10.2988
1978	1.4535	0.2280	6.3760	1978	0.6730	0.0593	11.3457
1979	0.9384	0.1691	5.5487	1979	0.8248	0.0729	11.3144
1980	1.9290	0.2945	6.5494	1980	0.7478	0.0716	10.4388
1981	2.6083	0.4449	5.8632	1981	0.8324	0.0798	10.4350
1982	1.3878	0.2230	6.2239	1982	0.9437	0.0897	10.5222
1983	1.4587	0.2343	6.2266	1983	1.0689	0.1024	10.4353
1984	1.4792	0.2415	6.1253	1984	1.1807	0.1146	10.3041
1985	1.3315	0.2239	5.9462	1985	1.1323	0.1130	10.0207
1986	1.8749	0.3001	6.2467	1986	1.0769	0.1104	9.7499
1987	2.0346	0.3188	6.3823	1987	1.5163	0.1489	10.1831
1988	2.1618	0.3400	6.3578	1988	1.6168	0.1586	10.1938
1989	1.8534	0.2968	6.2437	1989	1.7571	0.1710	10.2761
1990	1.6008	0.2676	5.9820	1990	1.6598	0.1634	10.1601
1991	2.0984	0.3301	6.3566	1991	1.7131	0.1681	10.1919
1992	1.6743	0.2727	6.1395	1992	1.6147	0.1581	10.2100
1993	1.6770	0.2773	6.0480	1993	1.8178	0.1763	10.3114
1994	1.5012	0.2517	5.9640	1994	2.1026	0.2003	10.4967
1995	1.8323	0.3012	6.0827	1995	2.9176	0.2732	10.6789
1996	1.7361	0.2781	6.2435	1996	2.4759	0.2330	10.6254
1997	1.3738	0.2337	5.8782	1997	2.0796	0.1968	10.5672
1998	1.7816	0.2873	6.2005	1998	2.1619	0.2051	10.5404
1999	1.6650	0.2741	6.0749	1999	1.8091	0.1757	10.2970
2000	2.0762	0.3330	6.2346	2000	1.2023	0.1270	9.4708
2001	2.0762			2001	1.5848	0.1776	8.9224

Note: The null hypotheses for no time effects are $H_0 : \pi_{1976}^2 = \cdots = \pi_{2000}^2 = 1$ and $H_0 : \phi_{1976}^2 = \cdots = \phi_{2001}^2 = 1$. The corresponding $\chi^2(25)$ and $\chi^2(26)$ test statistics equal 1370 (p-value=0.000) and 13819 (p-value=0.000) respectively.

Table 3: Estimated parameters of time effects

r_b^2	Estimate	SE	t-value	r_b^2	Estimate	SE	t-value
1916-1924	1.0000			1952	0.9080	0.3844	2.3619
1925	0.8654	0.3416	2.5338	1953	0.6905	0.2982	2.3157
1926	0.7315	0.3654	2.0021	1954	0.5926	0.2683	2.2085
1927	0.5494	0.3859	1.4238	1955	0.5762	0.2828	2.0374
1928	0.9315	0.3400	2.7394	1956	0.6421	0.3139	2.0454
1929	1.0717	0.3789	2.8285	1957	0.5738	0.2817	2.0369
1930	0.8019	0.3346	2.3965	1958	0.6288	0.3103	2.0265
1931	0.9997	0.3385	2.9531	1959	0.5730	0.2849	2.0110
1932	1.3688	0.4761	2.8752	1960	0.7619	0.3756	2.0288
1933	1.4552	0.5123	2.8403	1961	0.8936	0.4357	2.0507
1934	1.1025	0.3758	2.9339	1962	1.0785	0.5261	2.0500
1935	1.1671	0.3911	2.9846	1963	1.1581	0.5642	2.0526
1936	1.1062	0.3709	2.9823	1964	1.2306	0.5998	2.0517
1937	0.9423	0.3159	2.9833	1965	1.2372	0.6042	2.0477
1938	1.2243	0.4159	2.9436	1966	1.1442	0.5630	2.0322
1939	1.3125	0.4611	2.8464	1967	1.4720	0.7212	2.0409
1940	1.6218	0.6222	2.6065	1968	1.4089	0.6879	2.0480
1941	1.4829	0.5472	2.7101	1969	1.5886	0.7757	2.0479
1942	1.3912	0.5252	2.6489	1970	1.2635	0.6177	2.0454
1943	1.5567	0.5931	2.6245	1971	1.5743	0.7695	2.0458
1944	1.5683	0.5932	2.6437	1972	1.4642	0.7160	2.0450
1945	1.2451	0.4522	2.7537	1973	1.2183	0.5995	2.0322
1946	1.3725	0.5252	2.6131	1974	1.5192	0.7532	2.0169
1947	1.2396	0.4730	2.6207	1975	1.4305	0.7085	2.0190
1948	1.1750	0.4555	2.5797	1976	1.4346	0.7072	2.0285
1949	1.0509	0.4075	2.5789	1977	1.5775	0.7803	2.0217
1950	1.0709	0.4290	2.4966	1978-1980	1.8947	0.9378	2.0203
1951	0.9949	0.4059	2.4515				

Note: The null hypothesis for no cohort effects in permanent component is $H_0 : r^2_{1925} = \cdots = r^2_{1978} = 1$. The corresponding $\chi^2(54)$ test statistic equals 18519 (p-value=0.000).

Table 4: Estimated cohort effects for permanent component

s_b^2	Estimate	SE	t-value	s_b^2	Estimate	SE	t-value
1916-1924	1.000			1952	1.863	0.171	10.917
1925	1.463	0.179	8.171	1953	1.922	0.176	10.930
1926	1.148	0.147	7.791	1954	1.982	0.181	10.955
1927	1.305	0.154	8.497	1955	2.052	0.189	10.871
1928	1.215	0.138	8.800	1956	2.072	0.192	10.795
1929	1.256	0.146	8.592	1957	1.954	0.181	10.818
1930	1.395	0.157	8.905	1958	2.117	0.199	10.652
1931	1.403	0.150	9.323	1959	2.081	0.193	10.767
1932	1.177	0.135	8.745	1960	2.356	0.222	10.611
1933	1.158	0.126	9.160	1961	2.156	0.203	10.610
1934	1.448	0.163	8.902	1962	2.098	0.195	10.737
1935	1.415	0.152	9.315	1963	2.213	0.205	10.784
1936	1.382	0.149	9.282	1964	2.147	0.201	10.705
1937	1.394	0.143	9.721	1965	2.319	0.214	10.856
1938	1.411	0.142	9.908	1966	2.354	0.222	10.587
1939	1.380	0.144	9.587	1967	2.259	0.214	10.542
1940	1.283	0.136	9.410	1968	2.570	0.249	10.337
1941	1.463	0.147	9.971	1969	2.315	0.224	10.320
1942	1.373	0.143	9.582	1970	2.212	0.212	10.447
1943	1.478	0.147	10.050	1971	2.403	0.237	10.152
1944	1.643	0.156	10.524	1972	2.313	0.226	10.240
1945	1.690	0.159	10.609	1973	2.647	0.257	10.306
1946	1.655	0.159	10.417	1974	2.778	0.280	9.919
1947	1.678	0.154	10.870	1975	2.866	0.299	9.594
1948	1.676	0.155	10.795	1976	2.760	0.290	9.514
1949	1.795	0.164	10.946	1977	2.597	0.317	8.181
1950	1.835	0.169	10.854	1978-1980	2.002	0.303	6.600
1951	1.936	0.180	10.761				

Note: The null hypothesis for no cohort effects in transitory components is $H_0 : s^2_{1925} = \cdots = s^2_{1978} = 1$. The corresponding $\chi^2(54)$ test statistic equals 24900 (p-value=0.000).

Table 5: Cohort effects for transitory components

References

Abowd, John M and David Card (1989). On the covariance structure of earnings and hours changes. *Econometrica 57*(2), 411–45.

Altonji, Joseph G and Lewis M Segal (1996). Small-sample bias in gmm estimation of covariance structures. *Journal of Business & Economic Statistics 14*(3), 353–66.

Attanasio, Orazio P. and Guglielmo Weber (2010). Consumption and saving: Models of intertemporal allocation and their implications for public policy. (15756).

Baker, Michael (1997). Growth-rate heterogeneity and the covariance structure of life-cycle earnings. *Journal of Labor Economics 15*(2), 338–75.

Baker, Michael and Gary Solon (2003, April). Earnings dynamics and inequality among canadian men, 1976-1992: Evidence from longitudinal income tax records. *Journal of Labor Economics 21*(2), 267–288.

Beach, Charles, Ross Finnie, and David M. Gray (2010). Long-run inequality and annual instability of mens and womens earnings in canada. *Review of Income and Wealth 56*(3), 572–596.

Bird, D. (2004). Methodology for the 2004 annual survey of hours and earnings. *Labour Market Trends*, 457?64.

Blundell, Richard and Ian Preston (1995, August). Income, expenditure and the living standards of uk households. *Fiscal Studies 16*(3), 40–54.

Bound, John and Alan B. Krueger (1991). The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? Technical report.

Buchinsky, Moshe and Jennifer Hunt (1999). Wage mobility in the united states. *The Review of Economics and Statistics 81*(3), 351–368.

Chamberlain, G. (1984). Panel data. *Handbook of Econometrics*.

Constantinides, George M and Darrell Duffie (1996). Asset pricing with heterogeneous consumers. *Journal of Political Economy 104*(2), 219–40.

Deaton, Angus and Christina Paxson (1994). Intertemporal choice and inequality. *Journal of Political Economy 102*(3), 437–67.

Dickens, Richard (2000). The evolution of individual male earnings in great britain: 1975-95. *Economic Journal 110*(460), 27–49.

Doris, Aedin, Donal O'Neill, and Olive Sweetman (2010). Gmmcovearn: A stata module for gmm estimation of the covariance structure of earnings. (n212-10.pdf).

Gittleman, Maury and Mary Joyce (1994). Earnings mobility and long-run inequality: An analysis using matched cps data, type=Working paper,.

Gottschalk, Peter and Robert Moffitt (1994). The growth of earnings instability in the u.s. labor market. *Brookings Papers on Economic Activity 25*(2), 217–272.

Gottschalk, P. and R. Moffitt (1995). Trends in the covariance structure of earnings in the united states: 1969-1987. (1001-93).

Gottschalk, Peter and Enrico Spolaore (2000). On the evaluation of economic mobility.

Gottschalk, Peter T (1982). Earnings mobility: Permanent change or transitory fluctuations. *The Review of Economics and Statistics 64*(3), 450–56.

Guvenen, Fatih (2009). An empirical investigation of labor income processes. *Review of Economic Dynamics 12*(1), 58–79.

Haider, Steven J (2001). Earnings instability and earnings inequality of males in the united states: 1967-1991. *Journal of Labor Economics 19*(4), 799–836.

Hall, Robert E. (1971). The measurement of quality change from vintage price data. *Price Indexes and Quality Change*.

Huggett, Mark (1996, December). Wealth distribution in life-cycle economies. *Journal of Monetary Economics 38*(3), 469–494.

Kalwij, Adriaan S. and Rob Alessie (2007). Permanent and transitory wages of british men, 1975-2001: year, age and cohort effects. *Journal of Applied Econometrics 22*(6), 1063–1093.

Lillard, Lee A and Robert J Willis (1978). Dynamic aspects of earning mobility. *Econometrica 46*(5), 985–1012.

Lucas, Robert E. (2003). Macroeconomic priorities. *American Economic Review 93*(1), 1–14.

MaCurdy, Thomas E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of Econometrics 18*(1), 83–114.

Ramos, Xavier (2003). The covariance structure of earnings in great britain, 1991-1999. *Economica 70*(278), 353–374.

Shin, Donggyun and Gary Solon (2011). Trends in men's earnings volatility: What does the panel study of income dynamics show? *Journal of Public Economics 95*(7-8), 973–982.

Shorrocks, A. (1978). the measure of mobility. *Econometrica 46*, 1013–1024.

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron (2004, June). Cyclical dynamics in idiosyncratic labor market risk. *Journal of Political Economy 112*(3), 695–717.