

Gustavsson, Magnus

Working Paper

Earnings Dynamics and Inequality during Macroeconomic Turbulence: Sweden 1991-1999

Working Paper, No. 2002:20

Provided in Cooperation with:

Department of Economics, Uppsala University

Suggested Citation: Gustavsson, Magnus (2002) : Earnings Dynamics and Inequality during Macroeconomic Turbulence: Sweden 1991-1999, Working Paper, No. 2002:20, Uppsala University, Department of Economics, Uppsala, <https://nbn-resolving.de/urn:nbn:se:uu:diva-4451>

This Version is available at:

<https://hdl.handle.net/10419/82873>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Earnings Dynamics and Inequality during Macroeconomic Turbulence: Sweden 1991-1999^{*}

Magnus Gustavsson

December 5, 2002

Abstract

Using a rich longitudinal database, I study the dynamics behind changes in the distribution of annual earnings in Sweden 1991 to 1999. The analysis indicates a systematic increase in persistent earnings differentials during the 1990s; workers with low relative earnings in the beginning of the decade generally have lower relative earnings in the end. The evolution of transitory earnings fluctuations differs between age groups and matters the most for changes in earnings dispersion among younger workers. To shed some light on the role of wages behind the findings for earnings, wage dynamics is also studied. The results indicate that increased persistent wage inequality is one source behind the increase in persistent earnings inequality.

JEL classification: C23, D31, J31.

Keywords: Earnings Dynamics, Earnings Inequality, Wage Structure.

^{*}I thank Per-Anders Edin, Patrik Hesselius, Bertil Holmlund, Mårten Palme, and seminar participants at Uppsala University for useful comments and suggestions. I also thank Helena Sjödin at Statistics Sweden for providing information about the LINDA database. Financial support from The National Institute of Working Life and Swedish Model in Transition is gratefully acknowledged.

Correspondence: Department of Economics, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden.
E-mail: Magnus.Gustavsson@nek.uu.se

1. Introduction

The 1990s in Sweden was characterised by macroeconomic turbulence. The first years of the decade witnessed the worst economic downturn since the 1930s; GDP fell three years in a row, 1991 to 1993, by 4.7 percent altogether and the unemployment rate increased from 1.6 percent 1990 to 8.2 percent 1993. The economy remained in a recession until around 1997 when a noticeable recovery began; by 1999, GDP-growth was as high as 4.1 percent and the unemployment rate had decreased to 5.6 percent.¹

Earnings dispersion has markedly changed during this period. Aaberge *et al* (2000) find that annual earnings dispersion increased during the economic downturn in the beginning of the decade, and descriptive statistics in Edin & Fredriksson (2000) indicates a sharp lasting rise in annual earnings dispersion during the same period. Wage dispersion was approximately unchanged during the downturn, but markedly increased during the second half of the decade; see Statistics Sweden (2001b).²

The purpose of this paper is to study the dynamics behind changes in Swedish earnings dispersion during the 1990s. The analysis is based on the concept that earnings inequality can be decomposed into two separate components, one persistent and one transitory, and that the magnitude of these components may change over time. In this framework, persistent inequality originates from differences in education, training, and other systematic elements that determine a worker's relative earnings. The transitory component captures stochastic earnings differences that are independent of workers' systematic elements, and hence, that last for a shorter period.

¹ For a detailed picture of the labour market flows during the period at hand, see Holmlund & Vejsiu (2001). For information about GDP and other relevant statistics, see Statistics Sweden (2001a).

² For the evolution of wage inequality in Sweden prior to the 1990s, see Edin & Holmlund (1995).

The division of earnings inequality into its persistent and transitory components can help evaluate the role of individual skills in determining a worker's labour market outcome during the 1990s. Increased persistent inequality implies a deteriorated relative labour market position, in terms of wages and/or hours worked, for low skilled workers. In contrast, relative labour market positions will be unaltered by changes in transitory inequality.

The distinction between persistent and transitory inequality can also provide information about welfare changes. Specifically, an increase in persistent inequality implies a lasting increase in the inequality of consumption sets available to individuals, and this reduces welfare for most social welfare functions.³ On the other hand, an increase in transitory inequality will lead to increased year-to-year inequality of consumption sets, but there will be no change in consumption inequality when measured over several years. However, as optimising agents want to smooth consumption over transitory earnings fluctuations, increased fluctuations are by itself a potential source for decreased individual welfare. Only if one assumes that individuals have perfect foresight and can borrow and lend freely will earnings fluctuations be welfare neutral. But with more realistic assumptions, like uncertainty and the existence of liquidity constraints, increases in earnings instability decrease individual welfare.

Recently, several studies have forcefully demonstrated the empirical importance of decomposing changes in earnings inequality into its persistent and transitory components. Haider (2001) and Moffitt & Gottschalk (2002) use individual longitudinal data to study the dynamics behind changes in male earning dispersion in the U.S. from the late 1960s throughout the 1980s. Although their results differ somewhat, both studies find that persistent

³ Of course, changes in taxes and benefits might in some extreme case be able to neutralise increased earnings inequality; this is not the case for Sweden during the 1990s. Also, Lindbeck *et al* (1993) claim that wage and earnings differentials in Sweden are too small due to the union emphasis on egalitarian principles in the centralized wage bargaining present until mid 1980s. The long-term welfare effects from a systematic increase in earnings differentials in Sweden may therefore be somewhat blurrier than that for many other countries.

and transitory inequality contributed fairly equal to the observed increase in earnings dispersion from the late 1970s through the 1980s. Similar results, with a fairly equal contribution from the two components behind the increase in male earnings dispersion, are also reported for Canada 1976 to 1992 and for the U.K. 1975 to 1995 in Baker & Solon (1999) and in Dickens (2000), respectively.

In this paper I follow the recent literature and use a rich Swedish longitudinal database to decompose changes in male earnings dispersion 1991 to 1999 into its persistent and transitory components. The results indicate that persistent earnings inequality increased quite dramatically during the first years of the 1990s and then remained at the new higher level throughout the decade. The evolution of transitory earnings fluctuations differs between age groups and matters the most for earnings dispersion among younger workers. However, all age groups show a pattern of lower earnings fluctuations in the middle of the decade, i.e. during the macroeconomic slump. To illuminate the role of wages behind the increase in persistent earnings inequality, I also decompose changes in wage dispersion into its persistent and transitory components using a sub-sample of the sample used for earnings. The results indicate that increased persistent wage inequality is one source behind the increase in persistent earnings inequality.

The rest of the paper is organised as follows. The next section describes the framework of persistent and transitory earnings inequality used throughout the analysis. Section 3 presents the data and the sample construction. As an aid in the specification of the econometric model of persistent and transitory earnings inequality, section 4 presents an overview of the covariance structure of earnings during the 1990s. Section 5 describes the econometric model and estimation method. Section 6 contains the results from the econometric analysis and their implications. The paper ends with concluding remarks.

2. A Framework for the Analysis

The importance of persistent and transitory components in earnings variation is closely related to the extent of earnings mobility, i.e. the rate at which individuals shift positions in the earnings distribution. Due to its recurring importance in various research areas, earnings mobility is the subject of a voluminous empirical literature.⁴ To explain the connection between the models in the literature and the models estimated in this paper, I will begin with the basic permanent earnings model introduced by Friedman & Kuznets (1954). This model also provides a clear representation of how the covariance structure of earnings is related to persistent and transitory earnings inequality.

Let Y_{ibt} denote the log of earnings in year t of the i^{th} individual born in year b . Then

$$(1) \quad Y_{ibt} = \mu_{bt} + y_{ibt}$$

expresses Y_{ibt} as the cohort specific mean μ_{bt} in year t plus an individual specific deviation y_{ibt} from that mean. The model for y_{ibt} is

$$(2) \quad y_{ibt} = \alpha_{ib} + v_{ibt},$$

where α_{ib} is the persistent, or permanent, component and v_{ibt} is the transitory component.

The persistent component reflects the personal qualities of the individual, education and training, location, and other systematic elements. Its population variance is assumed to be σ_{α}^2 , i.e. constant over the life cycle and the same across cohorts. The transitory component captures the chance and other short lasting factors that influence relative earnings in a particular period. Transitory shocks are assumed to be uncorrelated with the persistent

⁴ For Sweden, see Palme (1995), Björklund & Palme (2002), and the references therein. For other countries, see e.g. the survey by Atkinson, Bourguignon & Morrisson (1992).

component, be serially uncorrelated, and have population variance σ_v^2 . With these assumptions, the variance of current relative earnings is

$$(3) \quad \text{Var}(y_{ibt}) = \sigma_\alpha^2 + \sigma_v^2$$

and the auto-covariance of relative earnings between year t and $t - s$ is

$$(4) \quad \text{Cov}(y_{ibt}, y_{ib,t-s}) = \sigma_\alpha^2.$$

That is, the variance of current relative earnings exceeds the auto-covariance by σ_v^2 , the variance of transitory earnings. The auto-correlation of earnings between year t and year $t + s$ is

$$(5) \quad r(t, t + s) = \sigma_\alpha^2 / [\sigma_\alpha^2 + \sigma_v^2].$$

A large contribution of the persistent component therefore implies that workers' earnings are highly correlated over time, and hence, that mobility rates are low.

The basic permanent earnings model, although intuitive, does not allow for changes in earnings inequality over time. A simple way to incorporate this, as suggested by Katz (1994), is to use the enhanced model

$$(6) \quad y_{ibt} = p_t \alpha_{ib} + \lambda_t v_{ibt},$$

where p_t and λ_t are time-varying factor loadings on the persistent and transitory components of earnings. Here, the factor loading p_t can be thought of as reflecting the time-varying returns to skill. With this model, the variance of y_{ibt} is

$$(7) \quad \text{Var}(y_{ibt}) = p_t^2 \sigma_\alpha^2 + \lambda_t^2 \sigma_v^2.$$

An increase in either factor loading generates increased cross-sectional earnings inequality. The character of the change in inequality depends, however, on which factor loading that

changes. A persistent rise in p_t increases earnings inequality measured over several years as the relative labour market advantage of workers with chronically high earnings is enhanced. An increase in λ_t without any change in p_t generates increased cross-sectional earnings dispersion by raising year-to-year earnings volatility, but there is no change in earnings inequality measured over several years.

Changes in the factor loadings are directly related to changes in earnings mobility. Increases in p_t without a change in λ_t will cause the auto-covariances to grow in greater proportion than the variances, so auto-correlations increase. In contrast, increases in λ_t without a change in p_t will only increase the variances, so autocorrelations decrease. Equal proportional increases in the two factor loadings will leave mobility rates unchanged even though individual earnings instability will be increased.

In empirical studies, several of the assumptions made above about the dynamics of relative earnings have been found to be overly restrictive. Specifically, persistent inequality should be allowed to vary with age and transitory shocks should be allowed to last for several periods. How this should be incorporated into the econometric model depends on the covariance structure of earnings in the data; this covariance structure is therefore investigated in section 4.

3. Data

The analysis is based on data from the register-based longitudinal database LINDA containing a sample of about 290 000 individuals constructed to be representative for the Swedish population in each year; see Edin & Fredriksson (2000). The database contains information from tax reports together with various individual characteristics.

The definition of earnings employed throughout the analysis is total earnings in real SEK at 1999 prices from all jobs held by an individual during a year.⁵ Due to its origin from tax reports, the earnings variable should be free of the measurement errors that are common in survey data. A potential drawback is that the contents of labour earnings in LINDA a particular year are contingent on the tax legislation in that year. However, no sharp change in the definition of earnings has occurred during the period studied here.⁶

The target group in the sample selection is males born in Sweden aged 26 to 56. Like previous studies, the focus is on males to avoid problems with more irregular female labour market participation e.g. due to parental leave. The applied age range is to assure that included individuals are old enough to have completed their education and too young to be considered for early retirement.

In constructing the analysis sample, I apply the revolving balanced panel design developed by Baker & Solon (1999). First, I identify the sixteen two-year birth cohorts who are between the ages 26 to 56 for at least four years in the period 1991 to 1999. I then select all males who had positive earnings in each year that the age requirement is met, i.e. an individual must have consecutive years of positive earnings during his cohort's whole sample period in order to be included. This means that I lose individuals with particular long spells of joblessness. Note, however, that individuals in the sample can have joblessness spells up to 22 months long, depending on how they fall across the calendar year. Next, I drop individuals altogether from the sample who in any of their cohort's sample years received any form of pension, had earnings from active ownership of a firm, or lived outside of Sweden.⁷ To remove noise in the

⁵ Sickness benefits are included in the earnings variable. The conversion into 1999 SEK is done by the use of the Swedish Consumer Price Index (KPI).

⁶ A sharp change from past definitions occurred in 1991 due to "the tax reform of the century"; see Agell, Englund & Södersten (1996).

⁷ I do not exclude individuals who have participated in some form of education since I do not have information on whether it is traditional education or part of some form of labour market program.

data, I also drop individuals altogether who in any of their cohort's sample years had earnings below 8736 real SEK at 1999 prices. In appendix B it is shown that this earnings threshold changes some results, but that the main conclusions remain unaffected.⁸

The end result of the sample selection procedure is a balanced earnings panel for each cohort, with the panel length varying across cohorts. The pooled analysis sample contains 37,836 individuals and the sample size *for each cohort* matches the pooled sample sizes available in most other longitudinal datasets; see Table 1. The variance of log earnings for the pooled sample displays the expected pattern with a sharp increase up until 1994, a decrease in 1995, and is then quite constant for the rest of the sample period; see appendix A. Overall, the variance is higher in 1999 than in 1991.

A fully balanced panel design is not suitable for the current purpose because time and average age will be perfectly collinear, and it will be hard to separate the effects of time and age on earnings inequality. The revolving balanced panel, however, contains stacked balanced cohort panels observed for different years, and this breaks the direct link between average age and time.

An alternative approach would be to use an unbalanced panel design where any years of positive earnings are included. The advantage with this approach is that the resulting panel is more representative of individuals with positive earnings during a specific year. However, the problem is that various sample moments for a given cohort are based on different samples. For example, the variance in year t is based on all individuals with positive earnings in that year, but the auto-covariance between year t and year $t - s$ is based only on those individuals

⁸ The threshold corresponds to the lower limit for being forced to manually declare labour earnings in 1999; the results are not sensitive to changes in the threshold. A similar threshold fixed over time is found in Baker & Solon (1999), Dickens (2000), and Haider (2001). I have also experimented with a threshold that is 8736 real SEK at 1999 prices and then adjusted for the year-to-year changes in real wages during the decade. This causes approximately no changes in the estimation results

Table 1: Cohorts Included in the Revolving Balanced Panel

Birth year	Sample size	Years observed	Age in Initial Year
1939/40	1925	1991-95	52
1941/42	2173	1991-97	50
1943/44	2502	1991-99	48
1945/46	2617	1991-99	46
1947/48	2557	1991-99	44
1949/50	2425	1991-99	42
1951/52	2312	1991-99	40
1953/54	2268	1991-99	38
1955/56	2333	1991-99	36
1957/58	2161	1991-99	34
1959/60	2210	1991-99	32
1961/62	2139	1991-99	30
1963/64	2372	1991-99	28
1965/66	2604	1992-99	27
1967/68	2637	1994-99	27
1969/70	2601	1996-99	27

Note: Age is defined by the older of the birth cohorts in each two-year cohort.

with positive earnings in both these years. This makes it is hard to keep a clear interpretation of the covariance matrix in terms of persistent and transitory inequality. Furthermore, the econometric methodology used in this paper is based on how the covariance structure of earnings changes over time. It may therefore be difficult to say whether estimated changes over time is due to true time and age effects or due to the different sample compositions from year-to-year.⁹

Sample composition effects could also be present in my revolving balanced panel due to the focus on individuals with at least four years of consecutive earnings and because earnings covariances of different orders are observable for different number of cohorts who, in turn, might face different selection criteria. In appendix A, I find that the revolving balanced panel design exclude some of the individuals in the lower part of the earnings distribution, and that

⁹ Solon, Barsky & Parker (1994) find evidence of important compositional changes of the labour force over the business cycle in the U.S. I am not aware of any study for Sweden on this issue.

this leads to underestimates of the variance of log earnings. However, I also find that the revolving balanced panel gives an accurate picture of how the variance of log earnings changes over time.¹⁰ This is important since the primary focus in this paper is on changes in earnings inequality during the 1990s, rather than its absolute level. At the very least, the panel should provide a useful depiction of how earnings inequality has changed among those men with a relative continuous attachment to the labour market.

4. The Covariance Structure of Earnings

As is clear from equation (3) and (4), the covariance structure of earnings is closely related to the division of persistent and transitory inequality, and can therefore aid in the specification of the econometric model. To obtain this covariance structure, I first estimate μ_{bt} in equation (1) with the sample mean log of earnings for the two-year birth cohort b in year t . Deviations of observed log earnings Y_{ibt} from that mean is then treated as a measure of y_{ibt} , i.e. current relative earnings.¹¹ This simple procedure adjusts for age, year, and cohort effects on average earnings. Next, for each of the 16 sample cohorts, born 1939/40 through 1969/70, I construct the sample auto-covariance matrix of y_{ibt} .¹² These are 9x9 matrices for the 11 cohorts observed for the whole sample period and smaller dimension for the other cohorts. For all cohorts taken together, this creates a total of 605 distinct auto-covariance elements.

Figure 1 presents the variances and auto-covariances of y_{ibt} of lags 1, 3 and 5 years for the cohorts born 1943/44, 1949/50, 1955/56, 1961/62, and 1965/66.¹³ For all cohorts the auto-covariances are positive, which is consistent with the presence of a persistent individual

¹⁰ Baker & Solon (1999) reach the same conclusion about their revolving balanced panel.

¹¹ The same approach is also found in Baker & Solon (1999).

¹² The covariance matrices are estimated in the same way as in Abowd & Card (1989).

¹³ Information on all cohorts is available on request.

component of earnings. The auto-covariances decrease and become closer in magnitude at longer lags lengths; this indicates that the transitory component is serially correlated over time. A serially correlated transitory component has also been found in most previous studies; see e.g. Atkinson, Bourguignon & Morrisson (1992).

Compared to the basic framework in section 2, the serially correlated transitory component alters the interpretation of the covariance structure of earnings. With uncorrelated transitory shocks, transitory inequality will only be present in the variance of log earnings. But with serially correlated shocks, transitory inequality will also be presented in the auto-covariances of log earnings, although at a diminishing rate by the lag length of the auto-covariance. The shorter lag auto-covariances thus reflect both the persistent and transitory component of earnings inequality, while the longer lag auto-covariances largely reflect the persistent component.

To see how persistent and transitory inequality varies with age, Figure 2 displays the variances and selected auto-covariances by age for every second year beginning 1991. As the longer lag auto-covariances largely reflect the persistent component of earnings inequality, Figure 2 indicates that persistent inequality increases at a diminishing rate until around age 40 and then is quite stable. The difference between the variances and the longer lag auto-covariances reflects how the variance of the transitory component of relative earnings varies with age. The U-shaped profile of the variances versus the concave increasing pattern of the longer lag auto-covariances hence imply that the variance of the transitory component is a function of age or experience. A heteroskedastic transitory variance is also found for Sweden by Hause (1980) and for Canada by Baker & Solon (1999).

Figure 1: Auto-covariances of Log Earnings for Selected Cohorts

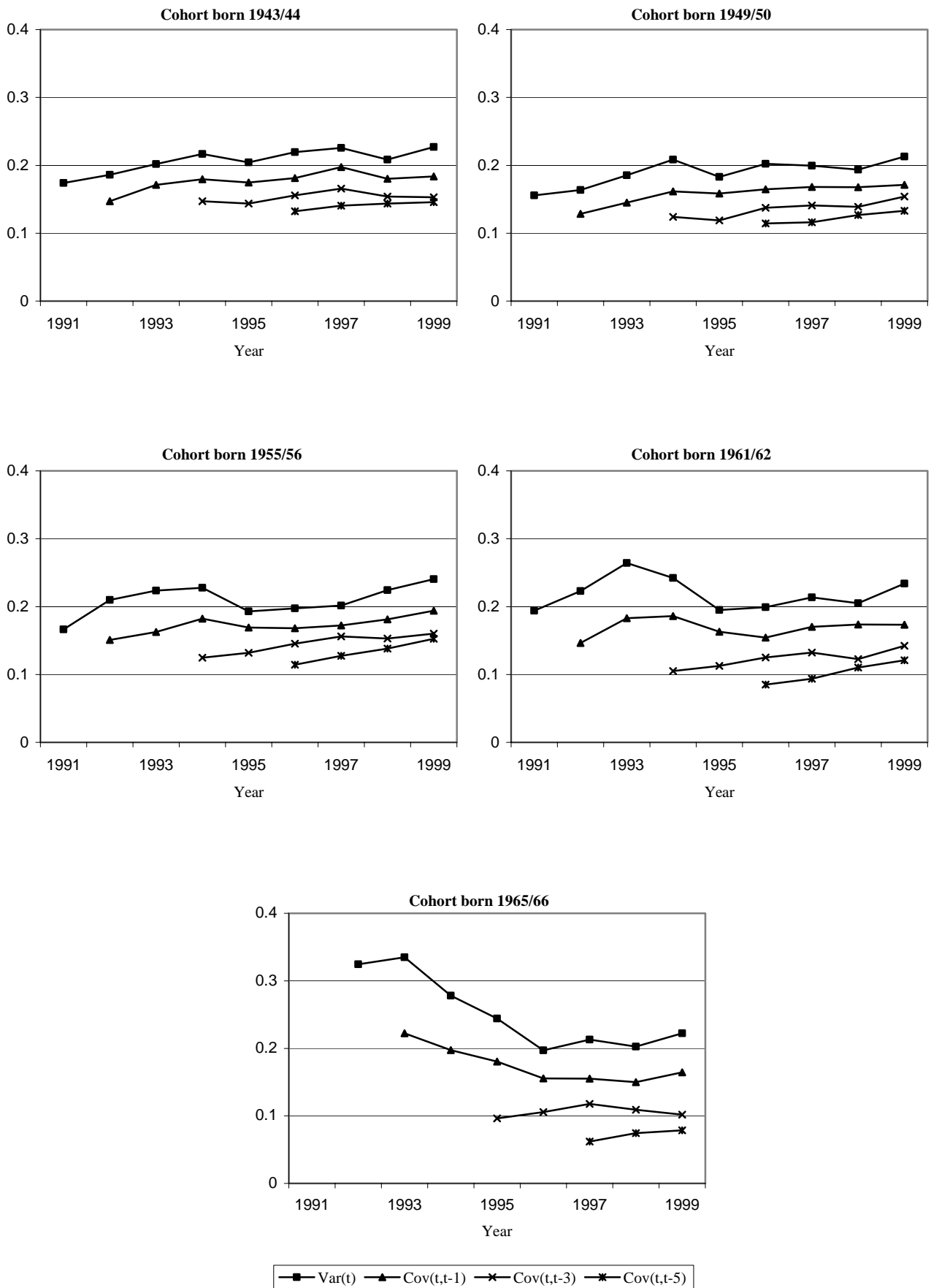
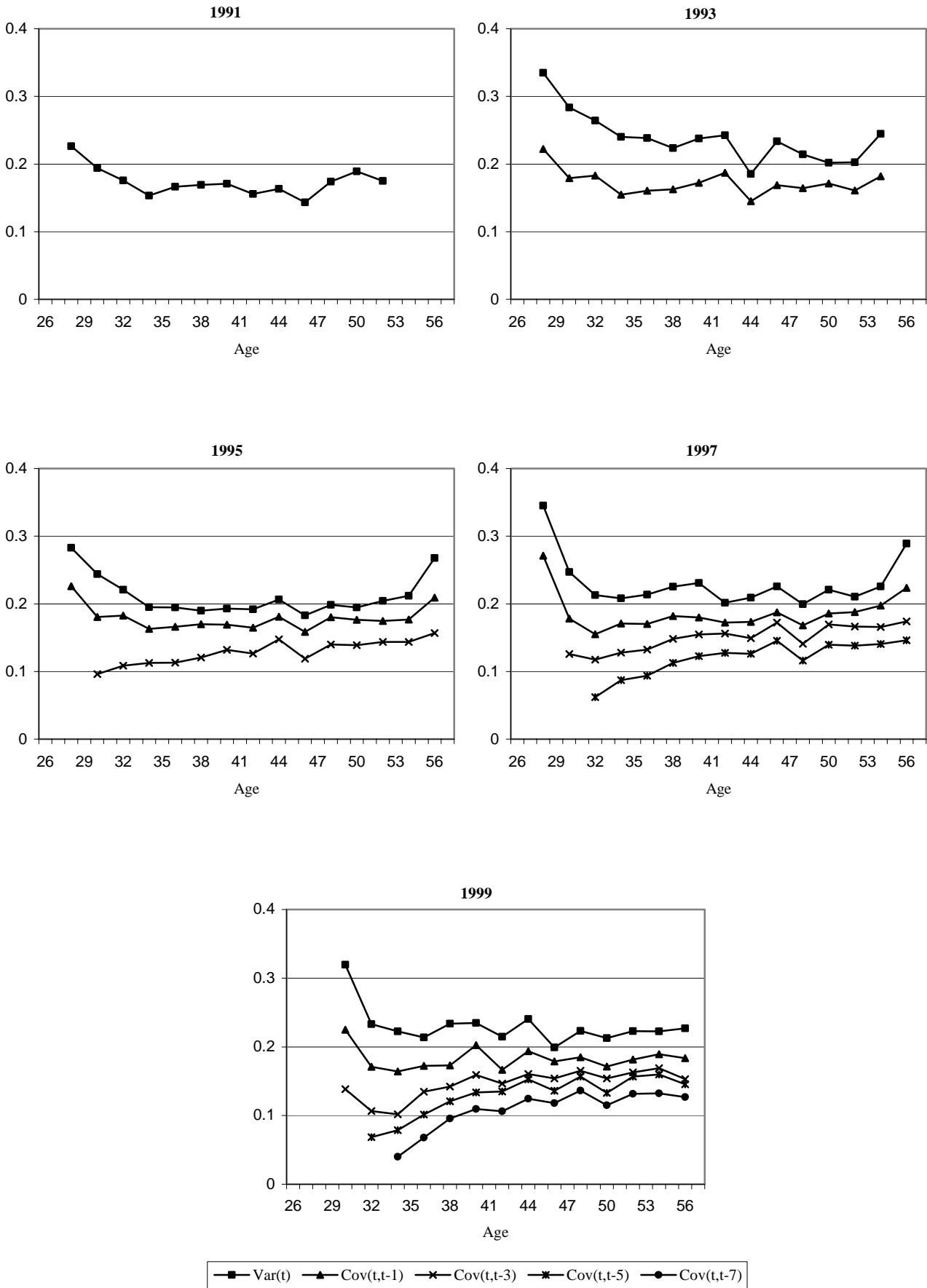


Figure 2: Auto-covariances of Log Earnings by Age



5. Econometric Model and Estimation Method

In specifying the econometric model, it is clear from section 4 that persistent inequality should be allowed to vary with age. The two main approaches in previous research to incorporate this have been to use either a “random growth” or a “random walk” specification in the persistent component.¹⁴

With the random growth specification, current relative earnings is

$$(8) \quad y_{ibt} = p_t (\delta_{ib} + \beta_{ib} a) + \lambda_t v_{ibt},$$

where $a = t - b - 27$, i.e. years since age 27 (the lowest observed age in the sample by definition). This specification incorporates individual-specific experience/earnings profiles: initial relative earnings, δ_{ib} , and earnings growth rates, β_{ib} , vary across individuals in a systematic manner.¹⁵ The economic models behind this specification include heterogeneity in human capital investments across equally able individuals, e.g. Becker (1975), or differences across jobs in the wage profiles for the purposes of effort extraction or bonding, i.e. Lazear (1979, 1981). With this model, the variance of y_{ibt} is

$$(9) \quad \text{Var}(y_{ibt}) = p_t^2 (\sigma_\delta^2 + 2a\sigma_{\delta\beta} + \sigma_\beta^2 a^2) + \lambda_t^2 \sigma_v^2,$$

where $\sigma_{\delta\beta}$ is the covariance between δ_{ib} and β_{ib} , which may be negative. Equation (9) implies that the longer lag auto-covariances of earnings, which mainly reflect persistent inequality, should display a convex quadratic pattern over the life cycle. As is clear from section 4, this is not consistent with the actual pattern in the data as the longer lag auto-covariances instead display a concave increasing pattern over the life cycle. I hence discard the random growth specification.

¹⁴ One exception is Baker & Solon (1999) who manage to incorporate both models.

¹⁵ Empirical evidence for the random growth model is found in, among other, Lillard & Weiss (1979), Hause (1980), and Haider (2001).

With the persistent component specified as a random walk in age, current relative earnings is

$$(10) \quad y_{ibt} = p_t u_{i,a} + \lambda_t v_{ibt}$$

where

$$(11) \quad u_{i,a} = u_{i,a-1} + r_{i,a}.$$

In equation (11), $r_{i,a} \sim \text{iid}(0, \sigma_{r,a}^2)$ is the innovation at each age.¹⁶ This specification is compatible with matching models where information about a worker's ability is revealed over time, e.g. Jovanovic (1979) and Farber & Gibbons (1996). With this model, the variance of y_{ibt} is

$$(12) \quad \text{Var}(y_{ibt}) = p_t^2 (\sigma_u^2 + \sum_a \sigma_{r,a}^2) + \lambda_t^2 \sigma_v^2,$$

where σ_u^2 is the variance of the first period shock (at age 27) in the persistent component.

Equation (12) implies that the longer lag auto-covariances of earnings never decrease with age. As the data at hand generally fulfil this implication, I use the random walk specification in the persistent component. In the estimations, to adjust this specification to the pattern in the data and save degrees of freedom, different restrictions are imposed on the innovation variances $\sigma_{r,a}^2$; these restrictions are discussed together with the results in the next section.

The evidence from section 4 together with previous studies also point to a serially correlated transitory component with a variance that vary with age. With these features incorporated, the complete model for current relative earnings is

$$(13) \quad y_{ibt} = p_t u_{ia} + \varepsilon_{ibt},$$

where

¹⁶ Empirical evidence for the random walk model is found in, among others, Topel (1991), Topel & Ward (1992), and Dickens (2000).

$$(14) \quad u_{ia} = u_{i,a-1} + r_{ia},$$

$$(15) \quad \varepsilon_{ibt} = \rho \varepsilon_{ib,t-1} + \lambda_t v_{ibt},$$

and

$$(16) \quad \text{Var}(v_{ibt}) = \gamma_0 + \gamma_1 a + \gamma_2 a^2 + \gamma_3 a^3 + \gamma_4 a^4.$$

Equation (15) allows the transitory component to follow a first order autoregressive process with year-specific factor loadings on the innovation v_{ibt} . This specification implies that luck follows luck in that a positive transitory shock dies out gradually. Furthermore, equation (16) allows the variance of v_{ibt} to be a quartic function of age.¹⁷

The parameters of the model in equation (13)-(16) are estimated by applying the minimum distance estimator of Chamberlain (1984), as described in Abowd & Card (1989).

Specifically, let \mathbf{C}_b contain the distinct elements of the population auto-covariance matrix of y_{ibt} for cohort b and let \mathbf{C} be an aggregate vector stacked with the \mathbf{C}_b vectors. Let the vector $\boldsymbol{\theta}$ contain all the parameters of the model and express the model's moment restrictions as $\mathbf{C} = f(\boldsymbol{\theta})$. The model in equation (13)-(16) then implies that the general variance element in \mathbf{C} is

$$(17) \quad \text{Var}(y_{ibt}) = p_t^2 (\sigma_u^2 + \sum_a \sigma_{ra}^2) + \rho^2 \text{Var}(\varepsilon_{ib,t-1}) + \lambda_t^2 (\gamma_0 + \gamma_1 a + \gamma_2 a^2 + \gamma_3 a^3 + \gamma_4 a^4),$$

and that the general auto-covariance element for year t and $t-s$ is

$$(18) \quad \text{Cov}(y_{ibt}, y_{ib,t-s}) = p_t p_{t-s} (\sigma_u^2 + \sum_{a-s} \sigma_{ra}^2) + \rho E[\varepsilon_{b,t-1} \varepsilon_{b,t-s}],$$

where $E[\cdot]$ is the expectation operator. The vector \mathbf{C} is estimated by the sample counterpart

$\hat{\mathbf{C}}$, and $\hat{\boldsymbol{\theta}}$ is chosen to minimize a distance function

¹⁷ Lower-order polynomials are generally rejected by the data.

$$(19) \mathbf{D} = (\hat{\mathbf{C}} - f(\hat{\boldsymbol{\theta}}))' \mathbf{W} (\hat{\mathbf{C}} - f(\hat{\boldsymbol{\theta}})),$$

where \mathbf{W} is a positive definite weighting matrix.

The asymptotically optimal choice of \mathbf{W} is the inverse of a matrix that consistently estimates the covariance matrix of \mathbf{C} . However, Altonji & Segall (1996) and Clark (1996) provide Monte Carlo evidence of potential serious finite sample bias in the estimate of $\boldsymbol{\theta}$ with this approach.¹⁸ I therefore follow the practise of the most recent literature and use the identity matrix as the weighting matrix. This “equally weighted minimum distance estimation” amounts to using non-linear least squares to fit $f(\hat{\boldsymbol{\theta}})$ to $\hat{\mathbf{C}}$. As outlined in Chamberlain (1984), the fourth moments of earnings in the sample is then used to estimate the standard errors of $\hat{\boldsymbol{\theta}}$.¹⁹

Finally, as can be seen in equation (17), cohort b 's overall variance in year t depends on its transitory variance in year $t - 1$. This raises the question of what the cohort's overall variance is in its first sample year. As pointed out by MaCurdy (1982), a time series approach to this problem is problematic since the assumption of infinite history is untenable. I therefore follow the approach of Baker & Solon (1999) and treat the initial transitory variances for the sixteen cohorts as sixteen additional parameters to be estimated. This approach recognises that earnings volatility varies across cohorts because they are at different stages of the life cycle and have lived through different times.

¹⁸ This bias arises because of correlated sampling errors in the second and fourth moments of earnings.

¹⁹ The standard errors are obtained from the following formula: $(\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{V}\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1}$, where \mathbf{G} is the gradient matrix $\partial f(\boldsymbol{\theta})/\partial \boldsymbol{\theta}$ evaluated at $\hat{\boldsymbol{\theta}}$ and \mathbf{V} is a block diagonal matrix containing the fourth moments of earnings for each cohort.

6. Results from the Econometric Analysis and their Implications

This section is divided into two parts. The first presents the results from the econometric analysis of changes in persistent and transitory earnings inequality. The second illuminates the role of wages behind changes in earnings inequality through a decomposition of wage inequality into its persistent and transitory components.

6.1 Earnings Dynamics

In the first estimated model of earnings dynamics, the innovation in the random walk is assumed to have two different variances: the first, $\sigma_{r,28-37}^2$, is valid up until age 37, the second, $\sigma_{r,38-56}^2$, is valid from age 38 to 56.²⁰ The first column of Table 2 presents the parameter estimates and associated standard errors for this model.

First in Table 2 are the variances of the persistent component, i.e. the variance of the first period shock (at age 27) and the two innovation variances. As the innovation variances are significantly larger than zero, the estimates indicate that persistent inequality increases over the whole life cycle. Also, the estimate of $\sigma_{r,28-37}^2$ is significantly larger than the estimate of $\sigma_{r,38-56}^2$, which indicates a faster growth in persistent inequality at younger ages.²¹

Next in Table 2 are the year-specific factor loadings on the persistent component. For identification, the value for 1991 is normalized to unity. The estimate for 1992 is significantly larger than 1 and the estimates for the following years are all significantly larger than that of

²⁰ Thus, the variance of the persistent component for individuals aged 56 in year t is $p_t^2(\sigma_u^2 + 10\sigma_{r,28-37}^2 + 19\sigma_{r,38-56}^2)$; the exact age-breakpoint does not matter for the final results of this model.

²¹ Wald tests are used throughout the analysis.

Table 2: Estimates of Earnings Dynamics Models

	(1)	(2)
Persistent Component		
σ_u^2	0.0469 (0.0035)	0.0367 (0.0036)
$\sigma_{r,28-37}^2$	0.0053 (0.0005)	
$\sigma_{r,38-56}^2$	0.00089 (0.00032)	
$\sigma_{r,38-52}^2$		
$\sigma_{r,28}^2$		0.0124 (0.0031)
$\sigma_{r,29-30}^2$		0.0082 (0.0016)
$\sigma_{r,31-32}^2$		0.0053 (0.0015)
$\sigma_{r,33-34}^2$		0.0043 (0.0016)
$\sigma_{r,35-36}^2$		0.0053 (0.0015)
$\sigma_{r,37-38}^2$		0.0026 (0.0015)
$\sigma_{r,39-40}^2$		0.0013 (0.0015)
$\sigma_{r,41-42}^2$		0.0003 (0.0016)
$\sigma_{r,43-44}^2$		-0.0007 (0.0014)
$\sigma_{r,45-46}^2$		0.0016 (0.0014)
$\sigma_{r,47-48}^2$		0.0020 (0.0015)
$\sigma_{r,49-50}^2$		0.0014 (0.0015)
$\sigma_{r,51-52}^2$		0.0005 (0.0016)
$\sigma_{r,53-54}^2$		0.0043 (0.0024)
$\sigma_{r,55-56}^2$		0.0118 (0.0060)
p_{91}	1.000	1.000
p_{92}	1.049 (0.013)	1.051 (0.013)
p_{93}	1.128 (0.016)	1.126 (0.016)
p_{94}	1.169 (0.019)	1.168 (0.019)
p_{95}	1.122 (0.019)	1.121 (0.020)
p_{96}	1.189 (0.021)	1.187 (0.022)
p_{97}	1.193 (0.021)	1.191 (0.022)
p_{98}	1.153 (0.021)	1.148 (0.021)
p_{99}	1.151 (0.020)	1.143 (0.020)
Transitory Component		
$\sigma_{1939/40}^2$	0.069 (0.017)	0.069 (0.018)
$\sigma_{1941/42}^2$	0.071 (0.016)	0.069 (0.016)
$\sigma_{1943/44}^2$	0.065 (0.014)	0.063 (0.014)
$\sigma_{1945/46}^2$	0.038 (0.011)	0.040 (0.012)
$\sigma_{1947/48}^2$	0.065 (0.014)	0.070 (0.014)
$\sigma_{1949/50}^2$	0.031 (0.014)	0.031 (0.014)
$\sigma_{1951/52}^2$	0.091 (0.016)	0.090 (0.016)
$\sigma_{1953/54}^2$	0.071 (0.017)	0.070 (0.017)
$\sigma_{1955/56}^2$	0.077 (0.015)	0.076 (0.016)
$\sigma_{1957/58}^2$	0.071 (0.015)	0.069 (0.015)
$\sigma_{1959/60}^2$	0.101 (0.017)	0.095 (0.017)
$\sigma_{1961/62}^2$	0.142 (0.019)	0.136 (0.019)
$\sigma_{1963/64}^2$	0.180 (0.017)	0.184 (0.017)

Table 2: (cont.)

	(1)	(2)
<i>Transitory Component</i>		
$\sigma_{1965/66}^2$	0.269 (0.018)	0.286 (0.018)
$\sigma_{1967/68}^2$	0.325 (0.021)	0.344 (0.021)
$\sigma_{1969/70}^2$	0.369 (0.023)	0.387 (0.023)
ρ	0.555 (0.011)	0.566 (0.010)
γ_0	0.201 (0.017)	0.189 (0.016)
γ_1	-0.034 (0.006)	-0.031 (0.006)
γ_2	0.003 (0.0008)	0.003 (0.0008)
γ_3	-1.35E-4 (3.63E-5)	-1.11E-4 (3.93E-5)
γ_4	2.11E-6 (5.86E-7)	1.60E-6 (6.74E-7)
λ_{92}	1.000	1.000
λ_{93}	1.048 (0.038)	1.054 (0.039)
λ_{94}	0.937 (0.042)	0.935 (0.043)
λ_{95}	0.854 (0.039)	0.846 (0.040)
λ_{96}	0.803 (0.044)	0.797 (0.046)
λ_{97}	0.945 (0.041)	0.946 (0.042)
λ_{98}	0.943 (0.040)	0.958 (0.042)
λ_{99}	1.030 (0.042)	1.056 (0.045)
SSE	0.055	0.051

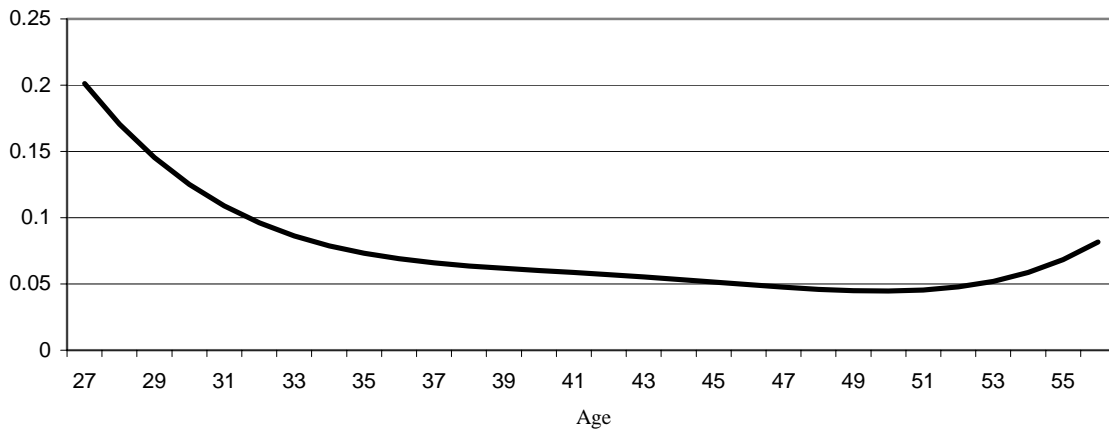
Note: Standard errors are in parentheses.

1992. The factor loadings thus indicate a lasting rise in persistent earnings inequality during the economic downturn.

The next section of Table 2 contains the parameters for the transitory component. First are the initial variances, which capture the accumulation of the transitory process up to the start of the sample period for each cohort. The estimates are all significantly larger than zero. Furthermore, Wald tests indicate that cohort differences in the initial variances are required. For example, a test of the hypothesis that the initial variances are the same for the cohorts born 1939/40 through 1957/58 ($H_0 : \sigma_{1939/40}^2 = \dots = \sigma_{1957/58}^2$) results in a p-value of 0.098 while the corresponding hypothesis for the younger cohorts ($H_0 : \sigma_{1959/60}^2 = \dots = \sigma_{1969/70}^2$) is strongly rejected at all normal levels.

The next block contains the autoregressive parameter ρ and the parameters of the quartic in age for the variance of the innovations to the transitory process. The estimate of ρ implies that more than 80 percent of a transitory shock has disappeared after 3 years. The estimated parameters of the innovation variance imply the U-shaped age profile displayed in Figure 3. As can be seen, the initial decline in the variance is quite sharp as it falls with over 60 percent to the mid 30's. Overall, the results of how transitory earnings fluctuations vary with age are consistent with the results in Baker & Solon (1999).²²

Figure 3: The Age Profile of the Variance of the Transitory Innovation



The final block of Table 2 reports the estimated year-specific factor loadings on the transitory innovations. Here, the factor loading for 1992 is normalised to unity as the innovation variance in 1991 is left unrestricted to identify the initial variances of the cohorts, i.e. the transitory variance in a cohort's first sample year is estimated solely by their initial variance. The estimated factor loadings are less precise compared to those on the persistent component, and only the estimates for 1995 and 1996 are significantly different (smaller) than 1.

²² None of the econometric models in Dickens (2000), Haider (2001), or Moffitt & Gottschalk (2002), allow for cohort differences in the initial transitory variance or for age differences in the variance of the innovations in the transitory component.

The second column of Table 2 presents a model with fewer restrictions on the persistent component; the variance of the innovations in the random walk are free to vary as of age 28 and after that only restricted to be the same for two ages at a time. The estimated factor loadings on both components and the parameters in the transitory process are approximately the same as in the previous column. The hypothesis that the variances of the random walk innovations are the same up to age 38 ($H_0 : \sigma_{r,28}^2 = \dots = \sigma_{r,37-38}^2$) cannot be rejected. The parameters for age 39-40 up to age 51-52 are imprecisely estimated and a test of the hypothesis that these are all equal to zero ($H_0 : \sigma_{r,39-40}^2 = \dots = \sigma_{r,51-52}^2 = 0$) cannot be rejected. One explanation for this result could be the close to zero variances for these ages in combination with the small number of observations on each age specific variance. The estimates of the last two parameters, for age 53-54 and 55-56, are both significantly larger than zero. One potential explanation for this result is increased differences in overall health at this stage in life. Another, not excluding explanation, is the possibility of larger cohort heterogeneity in skills among the older cohorts traceable back to a different education system.

The results in the second column of Table 2 indicate that the significantly larger than zero estimate of $\sigma_{r,38-56}^2$, in the first column of Table 2, might be an artefact of the noticeable increase in persistent inequality after age 52. However, this parameter is still significantly larger than zero if the innovation variances for individuals older than 52 years are left unrestricted in a model such as that in the first column of Table 2.²³

I have also estimated models with the random growth specification $\delta_{ib} + \beta_{ib}a$ in the persistent component. In all the results, the estimate of σ_{β}^2 is significantly *negative*. As outlined in section 5, this result is to be expected since the random growth model predicts a

²³ These results are available on request.

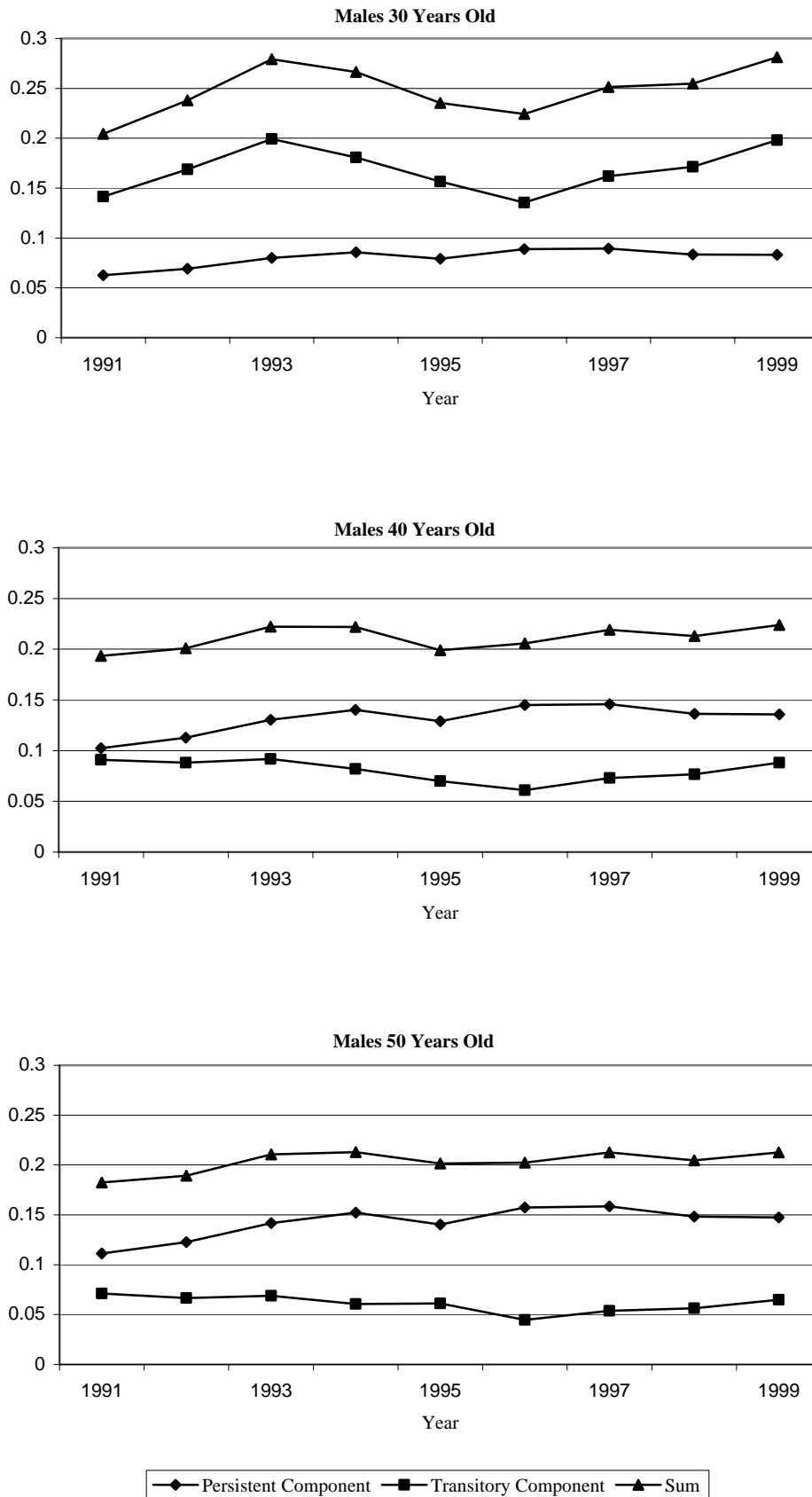
convex quadratic pattern of the longer lag auto-covariances over the life cycle, whereas the actual data at hand instead displays a concave increasing pattern. Furthermore, models where the transitory component follows an ARMA(1,1) process do not change any conclusions. In all the results, the moving average parameter is close to zero and insignificantly different from zero while the estimates of the other parameters are approximately the same as in Table 2.

As the results of the model in the first column of Table 2 are reasonable robust against different specifications, I use this model to decompose changes in earnings inequality into its persistent and transitory components. In studying the relative contributions of the two components, just plotting the time series of the factor loadings p_t and λ_t is not sufficient to give a full characterization of changes in inequality. The relative roles of persistent and transitory inequality depend not only on these two factor loadings but also on the relative magnitudes of the factors that they load, the initial transitory variances, and the autoregressive parameter. I therefore use the full model and predict the two components for each year holding age constant. Figure 4 displays these predictions performed for males 30, 40, and 50 years old. In moving from year to year, the factor loadings on the two components change, as does the initial variance used in generating the transitory variance up to these ages.²⁴

Figure 4, like the results in Table 2, make it clear that persistent inequality displays an upward trend during the 1990s. Another striking feature of Figure 4 is the different role played by the transitory component in the three age groups. Transitory earnings fluctuations dominate earnings dispersion among individuals 30 years old and are therefore mainly responsible for its changes during the sample period. For individuals 40 and 50 years old, transitory earnings inequality is lower than persistent inequality, but its evolution has still

²⁴ In fact, the initial variances changes every two years, corresponding to the cohort estimates in Table 2. For example, the initial variance for cohort 1951/52 is a direct estimate of the variance of the transitory component for individuals aged 40 in 1991. For 1993 I use the initial variance for cohort 1953/54, whose members are 40 in this year.

Figure 4: Decompositions of the Variance of Log Earnings



mattered. Figure 4 also shows that the evolution of transitory inequality differs between the age groups. Specifically, individuals 30 years old are the only ones who display a large increase until 1993. However, all age groups have a pattern of lower earnings fluctuations in the middle of the decade, i.e. during the macroeconomic slump.

The relative magnitude of the persistent and transitory components informs on the evolution of the earnings mobility rate, as defined in section 2. Since the evolution of transitory inequality differs between age groups, the evolution of the earnings mobility rate also differs. However, it is clear from Table 2 and Figure 4 that all age groups had the lowest mobility rate around 1996, since, in this year, persistent inequality was relative high while transitory inequality was relative low.

Overall, the results show that the 1990s in Sweden has been a decade of increased persistent inequality, while transitory inequality does not show any clear trend. Another finding is that transitory earnings inequality has been both larger and more volatile among younger individuals. As was discussed in the introduction, the social welfare consequences of larger persistent and transitory inequality are likely to be negative.²⁵ The increase in persistent inequality also means that workers with low relative earnings in the beginning of the decade generally have lower relative earnings in the end. As workers with low relative earnings generally also have lower skills, the results imply a deteriorated relative labour market position, in terms of wages and/or hours worked, for low skilled workers during the 1990s.

²⁵ However, the long-term welfare effects from a systematic increase in earnings differentials in Sweden may be somewhat blurrier than that for many other countries; see footnote 3.

6.2 The Role of Wages

Annual earnings are the product of hours worked and the hourly wage. It is important to know something about the relative contribution from these two factors in order to learn more about the labour market during the 1990s. This subsection therefore investigates whether persistent wage inequality also has increased.

Information in LINDA about individuals' monthly salaries in full time equivalents is available 1992 to 1999. The data is complete for the public sector but sampled in the private sector up to 1997, after which it is complete. The wage information is collected once a year and valid for the month of collection.

I use a balanced revolving panel with the same cohorts as in the analysis of earnings dynamics. Here, in addition to the requirements imposed for the earnings sample, consecutive years of positive wages are also required.²⁶ The samples for each cohort contain between 753 to 385 individuals and the average sample size is around 600 individuals. The total number of distinct auto-covariance elements for the pooled wage sample is 494. In appendix C, it is shown that the evolution of cross-sectional wage inequality in the balanced revolving panel is the same as in an unbalanced sample. However, the results in appendix C also point to the conclusion that information about wages in LINDA mainly is available for individuals with stable employment careers; this should be kept in memory when viewing the results below.

The covariance structure of wages speaks in favour of a specification where the persistent component follows a random walk in age. For the transitory component, I do not find evidence of serial correlation, neither of age differences in the innovation variance. The

²⁶ As the smaller sample sizes for each cohort increases the exposure to outliers, I remove individuals from the sample with monthly salaries below 6,000 or above 2,000,00 real SEK at 1999 prices.

covariance structure also reveals that the relative magnitude of the transitory component is much smaller than that for earnings. The small transitory wage component indicates that the transitory earnings component mainly consists of variation in hours worked. This can be compared to the results in Dickens (2000) and Haider (2001) that variation in wages make up a substantial part of transitory earnings in the U.K. and the U.S., respectively. However, part of my finding may be due to the use of somewhat different samples for earnings and wages.

Table 3 contains two specifications of wage dynamics that are consistent with the covariance structure of wages. The estimates in the first column indicate that persistent wage inequality, like persistent earnings inequality, increases significantly faster at younger ages. The year-specific factor loadings on the persistent component display an almost linear increase with estimates significantly different from 1 from 1995 and onward. The year-specific factor loadings on the transitory component are somewhat imprecisely estimated, but the estimates for 1998 and 1999 are significantly larger than 1.²⁷ As can be seen in the second column of Table 3, none of the results change with fewer restrictions on the variance of the random walk innovations. Overall, these results indicate that increased persistent wage inequality is one source behind the increase in persistent annual earnings inequality.

²⁷ One possible explanation for this is the greater coverage of the private sector in these years, presumed that transitory wage fluctuations are larger in the private sector. The greater coverage is due to the fact that movers from the public to the private sector in 1998 or 1999 are kept in the sample, whereas these movers tend to be excluded before 1998.

Table 3: Estimates of Wage Dynamics Models 1992-1999

	(1)	(2)
<i>Persistent Component</i>		
σ_u^2	0.0182 (0.0015)	0.0194 (0.0010)
$\sigma_{r,28-37}^2$	0.0035 (0.0003)	
$\sigma_{r,38-56}^2$	0.0014 (0.0002)	
$\sigma_{r,28}^2$		0.0031 (0.0007)
$\sigma_{r,29-30}^2$		0.0024 (0.0006)
$\sigma_{r,31-32}^2$		0.0041 (0.0008)
$\sigma_{r,33-34}^2$		0.0034 (0.0008)
$\sigma_{r,35-36}^2$		0.0035 (0.0010)
$\sigma_{r,37-38}^2$		0.0019 (0.0011)
$\sigma_{r,39-40}^2$		0.0022 (0.0010)
$\sigma_{r,41-42}^2$		0.0024 (0.0012)
$\sigma_{r,43-44}^2$		0.0017 (0.0012)
$\sigma_{r,45-46}^2$		0.0002 (0.0013)
$\sigma_{r,47-48}^2$		0.0017 (0.0012)
$\sigma_{r,49-50}^2$		0.0012 (0.0011)
$\sigma_{r,51-52}^2$		0.0013 (0.0010)
$\sigma_{r,53-54}^2$		0.0005 (0.0009)
$\sigma_{r,55-56}^2$		-0.0007 (0.0007)
p_{92}	1.000	1.000
p_{93}	1.010 (0.006)	1.010 (0.006)
p_{94}	1.014 (0.008)	1.016 (0.008)
p_{95}	1.030 (0.010)	1.033 (0.010)
p_{96}	1.035 (0.011)	1.038 (0.011)
p_{97}	1.058 (0.013)	1.062 (0.013)
p_{98}	1.064 (0.015)	1.065 (0.015)
p_{99}	1.074 (0.015)	1.076 (0.016)
<i>Transitory Component</i>		
γ_0	0.003 (0.0005)	0.003 (0.0005)
λ_{92}	1.000	1.000
λ_{94}	1.133 (0.094)	1.141 (0.087)
λ_{94}	1.198 (0.121)	1.210 (0.113)
λ_{95}	0.831 (0.094)	0.910 (0.089)
λ_{96}	1.059 (0.104)	1.053 (0.102)
λ_{97}	1.156 (0.120)	1.155 (0.117)
λ_{98}	1.408 (0.149)	1.422 (0.147)
λ_{99}	1.787 (0.173)	1.803 (0.175)
SSE	0.008	0.007

Note: Standard errors are in parentheses.

6. Concluding Remarks

This paper investigates the dynamics of annual male earnings dispersion in Sweden 1991 to 1999. Using a rich longitudinal database, I decompose earnings inequality into its persistent and transitory components, and study how these have changed over time. The results indicate a lasting rise in persistent inequality during the first years of the 1990s, a period characterised by Sweden's worst economic downturn in the post-war period. The evolution of transitory earnings fluctuations differs between age groups and matters the most for earnings dispersion among younger workers. However, all age groups show a pattern of lower earnings fluctuations in the middle of the decade, i.e. during the macroeconomic slump. The overall results also indicate that earnings mobility was at its lowest point around 1996 - just before the beginning of the macroeconomic recovery.

To illuminate the role of wages behind the increase in persistent earnings inequality, I also study wage dynamics from 1992 to 1999 using a sub-sample of the sample used for earnings. Compared to 1992, persistent wage inequality increased significantly from 1995 and onward. The results thus point to the conclusion that increased persistent wage inequality is one source behind the increase in persistent annual earnings inequality.

Several studies have found that higher unemployment is associated with lower relative earnings, and vice versa, among individuals in the lower part of the earnings distribution; see Björklund (1991). The estimated increase in persistent inequality during the first years of the 1990s is consistent with this mechanism. However, the finding that persistent inequality remained at a higher level, and even grew in some years, during the recovery in the second half of the decade is inconsistent with this as the sole explanation for the evolution of persistent inequality. Instead, other forces must have been at work besides this potential unemployment effect. One such force seems to be increased persistent wage

inequality. The cause for the lower transitory earnings fluctuations in the middle of the decade is neither fully understandable with today's knowledge of the Swedish labour market in the 1990s, although one could guess that the labour market was calmer during these years. To gain a better understanding of the driving forces behind the obtained results in this paper, and hence about the functioning of the labour market, more research is needed on how wages and hours worked have evolved among different earnings and age groups in Sweden during the 1990s.

References

- Aaberge, R, A Björklund, M Jäntti, P Pedersen, N Smith & T Wennemo (2000), “Unemployment Shocks and Income Distribution: How did the Nordic Countries Fare during their Crises?”, *Scandinavian Journal of Economics*, Vol. 102, pp. 77-99.
- Abowd, J & D Card (1989), “On the Covariance Structure of Earnings and Hours Changes”, *Econometrica*, Vol. 57, pp. 411-445.
- Agell, J, P Englund & J Södersten (1996), “Tax Reform of the Century – The Swedish Experiment”, *National Tax Journal*, Vol. 49, pp. 643-664.
- Altonji, J & L Segal (1996), “Small-Sample Bias in GMM Estimation of Covariance Structures”, *Journal of Business and Economic Statistics*, Vol. 14, pp. 353-366.
- Atkinsons, A, F Bourguignon & C Morrisson (1992), *Empirical Studies of Earnings Mobility*, Harwood Academic Publishers.
- Baker, M (1997), “Growth Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings”, *Journal of Labor Economics*, Vol. 15, pp. 338-375.
- Baker, M & G Solon (1999), “Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Tax Records”, Working Paper no. 7370, NBER.
- Becker, G (1975), *Human Capital*, University of Chicago Press.
- Björklund, A & M Palme (2002), “Income Redistribution Within the Life Cycle Versus Between Individuals: Empirical Evidence Using Swedish Panel Data”, forthcoming in Cohen, D, T Piketty & G Saint-Paul (eds.), *The New Economics of Rising Inequality*, Oxford University Press.
- Björklund, A (1991), “Unemployment and Income Distribution: Time-Series Evidence from Sweden”, *Scandinavian Journal of Economics*, Vol. 93, pp. 457-465.
- Chamberlain, G (1984), “Panel Data”, in Griliches, Z & M Intriligator (eds.), *Handbook of Econometrics*, Vol. 2, Elsevier Science B.V.
- Clark, T (1996), “Small-Sample Properties of Estimators of Nonlinear Models of Covariance Structure”, *Journal of Business and Economic Statistics*, Vol. 14, pp. 367-373.
- Dickens, R (2000), “The Evolution of Individual Male Earnings in Great Britain: 1975-95”, *Economic Journal*, Vol. 110, pp. 27-49.
- Edin, P-A & P Fredriksson (2000), “LINDA – Longitudinal Individual Data for Sweden” Working Paper 2000:19, Department of Economics, Uppsala University.

- Edin, P-A & B Holmlund (1995), "The Swedish Wage Structure: The Rise and Fall of Solidarity Wage Policy?", in Freeman, R & L Katz (eds.), *Differences and Changes in Wage Structures*, University of Chicago Press.
- Farber, H & R Gibbons (1996), "Learning and Wage Dynamics", *Quarterly Journal of Economics*, Vol. 111, pp. 1007-1047.
- Friedman, M & S Kuznets (1954), *Income from Independent Professional Practise*, NBER.
- Haider, J (2001), "Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991", *Journal of Labor Economics*, Vol. 19, pp. 799-836.
- Hause, J (1980), "The Fine Structure of Earnings and the On-the-Job Training Hypothesis", *Econometrica*, Vol. 48, pp. 1013-1029.
- Holmlund, B & A Vejsiu (2001), "Labor Market Flows in Boom and Slump: Sweden 1987-2000", in Vejsiu, A, *Economic Studies 56*, Uppsala University.
- Jovanovic, B (1979), "Job Matching and the Theory of Turnover", *Journal of Political Economy*, Vol. 87, pp. 972-990.
- Katz, L (1994), "The Growth of Earnings Instability in the U.S. Labor Market: Comments and Discussion", *Brookings Paper on Economic Activity*, 1994:2, pp. 255-261.
- Lazear, E (1979), "Why Is There Mandatory Retirement?", *Journal of Political Economy*, Vol. 87, pp. 1261-1285.
- Lazear, E (1981), "Agency, Earnings Profiles, Productivity, and Hours Restrictions", *American Economic Review*, Vol. 71, pp. 606-620.
- Lillard, L & Y Weiss (1979), "Components of Variation in Panel Earnings Data: American Scientists 1960-70", *Econometrica*, Vol. 47, pp. 437-454.
- Linbeck, A, P Molander, T Persson, O Peterson, A Sandmo, B Swedenborg & N Thygesen (1993), "Options for Economic and Political Reform in Sweden", *Economic Policy*, Vol 8, pp. 219-263.
- MaCurdy, T (1982), "The Use of Time Series Processes to Model the Error Structure of Earnings in Longitudinal Data Analysis", *Journal of Econometrics*, Vol. 18, pp. 83-114.
- Moffitt, R & P Gottschalk (2002), "Trends in the Transitory Variance of Earnings in the United States" *Economic Journal*, Vol. 112, pp. 68-73.
- Palme, M (1995), "Earnings Mobility and Distribution: Comparing Statistical Models on Swedish Data", *Labour Economics*, Vol. 2, pp. 213-247.
- Solon, G, R Barsky & J Parker (1994), "Measuring the Cyclicity of Real Wages: How Important is Composition Bias?", *Quarterly Journal of Economics*, Vol. 109, pp. 1-25.

Statistics Sweden (2001a), *Statistical Year Book of Sweden 2002*, SCB Örebro.

Statistics Sweden (2001b), *Statistical Year Book of Salaries and Wages 2000*, SCB Örebro.

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

Topel, R & M Ward (1992), "Job Mobility and the Careers of Young Men", *Quarterly Journal of Economics*, Vol. 107, pp. 439-479.

Appendix A: Earnings Inequality with Different Panel Designs

The revolving balanced panel design could lead to wrong conclusions about earnings inequality due to sample composition effects. To investigate this, the second to the fourth columns of Table A1 contain the age interval, sample size, and the variance of log earnings, for each of the sample years in the pooled revolving balanced panel. The next two columns of Table A1 contain an unbalanced sample with the same age intervals as the revolving panel, where all individuals with earnings larger than 8736 real SEK in any given year are included. The unbalanced sample contains about 20 to 30 percent more individuals, and most of the additional individuals are in the lower part of the earnings distribution; detailed descriptive statistics of earnings for these samples are available on request. As a consequence of the higher frequency of individuals with low earnings, the variances of log earnings in the unbalanced sample are markedly higher.

Table A1: Descriptive Statistics of Various Samples

Year	Revolving Panel Age Intervals	Revolving Panel		Unbalanced Sample, Revolving Panel Age Intervals		Unbalanced Sample, Age Interval 26-56	
		N	Var(Y)	N	Var(Y)	N	Var(Y)
1991	27-52	29,994	0.184	39,688	0.295	45,615	0.299
1992	26-53	32,598	0.227	42,681	0.361	45,693	0.356
1993	27-54	32,598	0.256	41,633	0.423	45,312	0.430
1994	26-55	35,235	0.269	44,192	0.431	45,210	0.429
1995	27-56	35,235	0.224	44,041	0.388	45,578	0.397
1996	26-55	35,911	0.246	44,401	0.402	45,418	0.401
1997	27-56	35,911	0.244	43,331	0.396	44,886	0.411
1998	28-55	33,738	0.227	40,701	0.370	45,121	0.394
1999	29-56	33,738	0.237	41,049	0.358	45,765	0.388

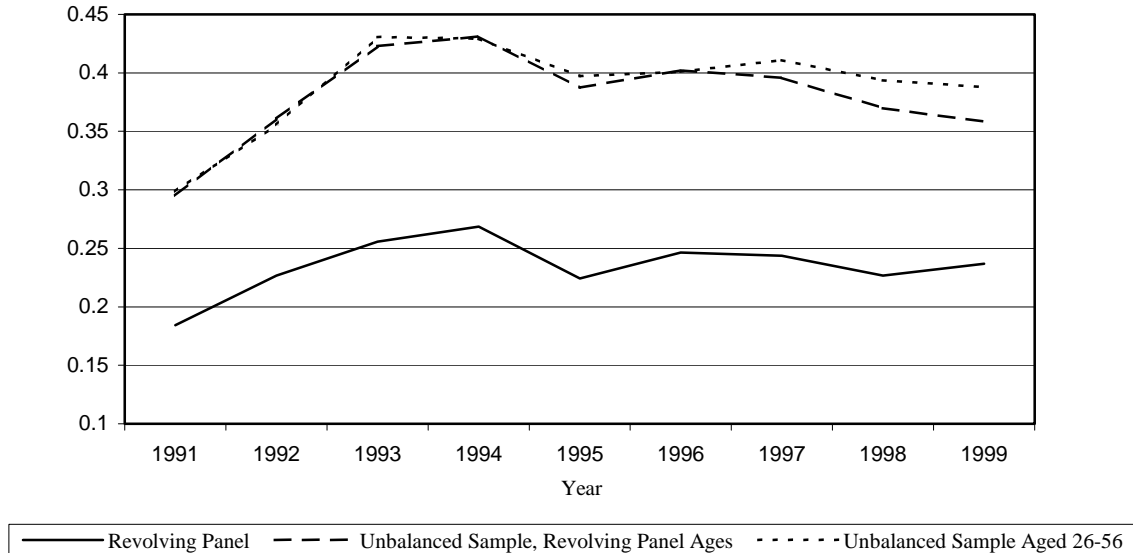
Note: Only individuals with earnings larger than 8736 real SEK at 1999 prices are included in any given year in the samples.

As my revolving balanced panel do age somewhat over time, I have also constructed an unbalanced sample with the age interval 26 to 56 in each year, i.e. a sample that do not age over time; the seventh and eighth columns of Table A1 describe this sample. Although there are some minor discrepancies between this sample and the sample in the previous two columns, it is clear that the requirement of consecutive years of earnings larger than 8736

real SEK, rather than marginal aging over time, accounts for the differences between this sample and my revolving balanced panel (setting the earnings threshold to zero in all samples leads to the same conclusion).

The evolution of the variance over time in the revolving panel and in the two unbalanced panels is also plotted in Figure A1. As can be seen, the revolving panel appears to shadow the variances in the two comparison panels. In fact, the correlation coefficient between the revolving panel and the unbalanced panel with the same age intervals and that with a fixed age interval is 0.95 and 0.93, respectively. This indicates that, though the revolving balanced panel design leads to smaller estimates of the variance of log earnings due to fewer individuals with low earnings, the evolution of the variance is quite unaffected by the panel design.

Figure A1: The Variance of Log Earnings in Various Samples



Appendix B: The Effect of the Earnings Threshold

Table A2 displays how the results for earnings dynamics change when the earnings threshold in the sample construction is set to zero. That is, when consecutive years of

earnings larger than zero is required instead of consecutive years of earnings larger than 8736 in real SEK at 1999 prices.

The estimated factor loadings on the persistent component in the first column of Table A2 follow the same pattern as in the main text, but their increase compared to 1991 is about twice as large. The variance of the first period shock in the persistent component, $\hat{\sigma}_u^2$, is larger than before while $\hat{\sigma}_{r28-37}^2$ is lower and $\hat{\sigma}_{r38-56}^2$ is not significantly larger than zero.

As the variance in the first period shock reflects the variance in persistent earnings across individuals as of age 27, the results indicate that the main part of persistent earnings inequality is determined before age 27. However, setting the earnings threshold to zero increases the sample size with about 40 to 50 individuals in each two-year birth cohort born 1939/40 to 1957/58, while the corresponding number for the younger cohorts are around 200 individuals. The removal of the threshold therefore results in proportionally larger estimates of the variances and auto-variances for the youngest cohorts. The large proportion of young individuals with periods of basically no attachment to the labour market, e.g. due to education, probably makes the data noisy, and this makes it hard to estimate the effect of age on persistent earnings inequality.

Next are the results for the transitory component. The estimated initial variances are more than twice as large for the three youngest cohorts, while the difference is smaller for older cohorts. The estimate of ρ is close to that from the main text while the estimated parameters in the innovation variance are larger in absolute value but display the same pattern as in Figure 3. The factor loadings display a similar pattern as in the main text, but with larger movements. The factor loading for 1996 is still significantly smaller than 1 while this no longer holds true for 1995. Another difference is the significantly larger than 1 factor loading for 1993.

Table A2: Estimates of Earnings Dynamics Models with the Earnings Threshold Equals Zero

	(1)	(2)
Persistent Component		
σ_u^2	0.0738 (0.0053)	0.0650 (0.0056)
$\sigma_{r,28-37}^2$	0.0031 (0.0007)	
$\sigma_{r,38-56}^2$	0.00017 (0.00041)	
$\sigma_{r,28}^2$		0.0119 (0.0056)
$\sigma_{r,29-30}^2$		0.0034 (0.0029)
$\sigma_{r,31-32}^2$		0.0018 (0.0028)
$\sigma_{r,33-34}^2$		0.0062 (0.0032)
$\sigma_{r,35-36}^2$		0.0029 (0.0030)
$\sigma_{r,37-38}^2$		-0.0022 (0.0029)
$\sigma_{r,39-40}^2$		0.0009 (0.0027)
$\sigma_{r,41-42}^2$		0.0005 (0.0030)
$\sigma_{r,43-44}^2$		0.0020 (0.0028)
$\sigma_{r,45-46}^2$		-0.0028 (0.0027)
$\sigma_{r,47-48}^2$		0.0036 (0.0027)
$\sigma_{r,49-50}^2$		-0.0040 (0.0025)
$\sigma_{r,51-52}^2$		0.0013 (0.0025)
$\sigma_{r,53-54}^2$		0.0072 (0.0044)
$\sigma_{r,55-56}^2$		0.0158 (0.0098)
p_{91}	1.000	1.000
p_{92}	1.105 (0.021)	1.113 (0.021)
p_{93}	1.249 (0.028)	1.256 (0.029)
p_{94}	1.390 (0.038)	1.401 (0.039)
p_{95}	1.286 (0.039)	1.296 (0.040)
p_{96}	1.407 (0.042)	1.417 (0.043)
p_{97}	1.419 (0.042)	1.427 (0.043)
p_{98}	1.334 (0.038)	1.330 (0.039)
p_{99}	1.334 (0.036)	1.323 (0.037)
Transitory Component		
$\sigma_{1939/40}^2$	0.091 (0.021)	0.097 (0.023)
$\sigma_{1941/42}^2$	0.145 (0.028)	0.155 (0.028)
$\sigma_{1943/44}^2$	0.063 (0.018)	0.057 (0.019)
$\sigma_{1945/46}^2$	0.064 (0.019)	0.067 (0.0196)
$\sigma_{1947/48}^2$	0.119 (0.029)	0.114 (0.030)
$\sigma_{1949/50}^2$	0.094 (0.026)	0.094 (0.026)
$\sigma_{1951/52}^2$	0.118 (0.022)	0.123 (0.023)
$\sigma_{1953/54}^2$	0.131 (0.032)	0.138 (0.032)
$\sigma_{1955/56}^2$	0.083 (0.023)	0.078 (0.024)
$\sigma_{1957/58}^2$	0.133 (0.033)	0.124 (0.033)
$\sigma_{1959/60}^2$	0.190 (0.036)	0.190 (0.037)
$\sigma_{1961/62}^2$	0.274 (0.042)	0.272 (0.042)
$\sigma_{1963/64}^2$	0.306 (0.032)	0.306 (0.032)
$\sigma_{1965/66}^2$	0.540 (0.047)	0.555 (0.047)

Table A2: (cont.)

	(1)	(2)
<i>Transitory Component</i>		
$\sigma_{1967/68}^2$	0.753 (0.056)	0.774 (0.055)
$\sigma_{1969/70}^2$	0.830 (0.060)	0.850 (0.059)
ρ	0.505 (0.012)	0.508 (0.0107)
γ_0	0.509 (0.044)	0.489 (0.043)
γ_1	-0.099 (0.015)	-0.091 (0.014)
γ_2	0.010 (0.0018)	0.009 (0.0019)
γ_3	-4.44E-4 (8.83E-5)	-3.62E-4 (9.27E-5)
γ_4	7.08E-6 (1.44E-6)	5.42E-6 (1.57E-6)
λ_{92}	1.000	1.000
λ_{93}	1.131 (0.047)	1.136 (0.047)
λ_{94}	0.950 (0.050)	0.942 (0.051)
λ_{95}	0.939 (0.047)	0.936 (0.049)
λ_{96}	0.891 (0.052)	0.885 (0.054)
λ_{97}	1.039 (0.051)	1.042 (0.052)
λ_{98}	0.954 (0.048)	0.969 (0.049)
λ_{99}	0.952 (0.043)	0.973 (0.045)
SSE	0.419	0.401

Note: Standard errors are in parentheses.

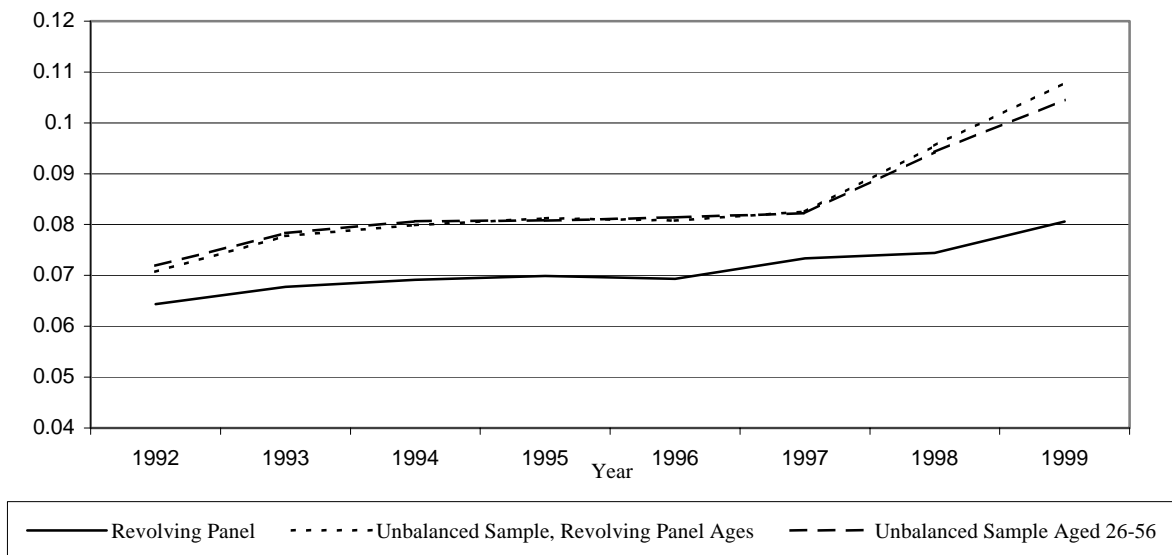
The second column of Table A2 contains the same model as in the second column of Table 2. For the variances in the random walk, only $\hat{\sigma}_u^2$, $\hat{\sigma}_{r28}^2$ and $\hat{\sigma}_{r33-34}^2$ are significantly larger than zero while $\hat{\sigma}_{r49-50}^2$ is on the verge of being significantly *negative*. The rest of the parameters in the persistent component are imprecisely estimated.

Appendix C: Sample Composition Effects in the Wage Sample

Figure A2 graphs the evolution of the variance of log wages in the revolving balanced panel, in an unbalanced panel with the same age intervals as the revolving panel, and in an unbalanced panel with fixed age intervals; detailed descriptive statistics for these samples are available on request. The variances of log wages are lower in the revolving balanced panel due to fewer individuals with low wages. As for earnings, it is also clear that marginal

aging over time in the revolving panel is of minor importance for the evolution of the variance of log wages. The larger wage dispersion in the private sector is clearly seen in the two unbalanced panels as wage dispersion makes a jump in 1998, the year when information on wages also becomes available for all individuals employed in the private sector. Despite this, the variances show similar movement and the correlations between the revolving panel and the unbalanced panels are around 0.96.

Figure A2: The Variance of Log Wages in Various Samples



Selection effects are probably present in the wage sample independent of panel design as the wage data is collected once a year and only valid for the month of collection. Individuals with stable employment careers thus have a higher probability of reporting a positive wage in any given year. An additional potential problem with the wage data is the overrepresentation of individuals working in the public sector.

One way to shed some light on the selection effects in the wage sample is to estimate a model of earnings dynamics for the wage sample, and then compare the results to those obtained from a sample that is based on all individuals. To obtain the comparison sample, I create a balanced revolving panel for earnings in the same manner as in the main text, with

the exception that I only focus on 1992 to 1999. This sample is henceforth referred to as the “complete sample”.

There are some problems with models of earnings dynamics based on the wage sample. Those specified as in the main text results in a negative estimate of some of the initial transitory variances. None of these problems are, however, present for the complete sample. The likely explanation is a large proportion of individuals with stable employment careers in the wage sample, which will cause the variance in the transitory part to be small and stable over time. The models thus become over-fitted in that irrelevant variables are included. Consequently, in Table A3 the initial variances are restricted to be the same for two two-year birth cohorts at a time except for the four youngest two-year birth cohorts. Also, three instead of five parameters are used for the innovation variance in the transitory component.

In Table A3, the factor loadings on the persistent component display an upward trend for both samples but the estimates for the complete sample fluctuate more. The pattern for the wage sample is also quite similar to that obtained for the estimates of wage dynamics in Table 3. The transitory part is somewhat imprecisely estimated for the wage sample, and comparisons with the complete sample is therefore hard. One exception is the estimate of the autoregressive parameter, which is precisely estimated and of the same magnitude in the two samples.

Overall, the results suggest that there are some problems with composition effects in the wage sample, probably due to overrepresentation of individuals with stable employment careers. One thing that speaks for this conclusion, as discussed above, is the difficulty of estimating the transitory earnings component for the wage sample. Another is the similar pattern of the factor loadings on the persistent component for earnings and wages. This similar pattern is probably due to small changes in persistent hours worked.

Table A3: Estimates of Earnings Dynamics 1992-1999

	Wage Sample (1)	Complete Sample 1992-99 (2)
<i>Persistent Component</i>		
σ_u^2	0.0325 (0.0034)	0.0540 (0.0036)
$\sigma_{r,28-37}^2$	0.0035 (0.0005)	0.0057 (0.0005)
$\sigma_{r,38-56}^2$	0.0015 (0.0004)	0.0011 (0.0003)
p_{92}	1.000	1.000
p_{93}	1.012 (0.010)	1.059 (0.009)
p_{94}	1.023 (0.014)	1.107 (0.012)
p_{95}	1.044 (0.015)	1.061 (0.012)
p_{96}	1.049 (0.015)	1.135 (0.013)
p_{97}	1.054 (0.016)	1.143 (0.013)
p_{98}	1.071 (0.017)	1.101 (0.013)
p_{99}	1.081 (0.018)	1.096 (0.012)
<i>Transitory Component</i>		
$\sigma_{1939-42}^2$	0.010 (0.008)	0.068 (0.009)
$\sigma_{1943-46}^2$	0.014 (0.006)	0.068 (0.007)
$\sigma_{1947-50}^2$	0.020 (0.008)	0.067 (0.009)
$\sigma_{1951-54}^2$	0.002 (0.006)	0.110 (0.007)
$\sigma_{1955-58}^2$	0.022 (0.011)	0.103 (0.010)
$\sigma_{1959-62}^2$	0.028 (0.012)	0.145 (0.012)
$\sigma_{1963/64}^2$	0.077 (0.021)	0.203 (0.020)
$\sigma_{1965/66}^2$	0.075 (0.023)	0.272 (0.023)
$\sigma_{1967/68}^2$	0.042 (0.013)	0.328 (0.013)
$\sigma_{1969/70}^2$	0.130 (0.022)	0.373 (0.021)
ρ	0.469 (0.046)	0.552 (0.010)
γ_0	0.021 (0.007)	0.186 (0.008)
γ_1	-0.0011 (0.0006)	-0.013 (0.0008)
γ_2	1.57E-5 (1.66E-5)	3.47E-4 (2.47E-5)
λ_{93}	1.000	1.000
λ_{94}	0.940 (0.390)	0.735 (0.045)
λ_{95}	0.580 (0.204)	0.639 (0.028)
λ_{96}	0.729 (0.333)	0.504 (0.040)
λ_{97}	0.789 (0.329)	0.741 (0.040)
λ_{98}	1.106 (0.422)	0.725 (0.041)
λ_{99}	1.535 (0.535)	0.864 (0.045)
SSE	0.017	0.061

Note: Standard errors are in parentheses