Noname manuscript No.
(will be inserted by the editor)

Earnings Dynamics of Men and Women in Finland: Permanent Inequality versus Earnings Instability

Otto Kässi

the date of receipt and acceptance should be inserted later

Abstract I decompose the earnings variance of Finnish male and female workers into its permanent and transitory components using the approach of Baker (1997) and Haider (2001) in the spirit of scientific replication.

I find that the increasing earnings inequality of men and women is driven by both the transitory and permanent components of earnings. In addition, I find considerable differences in the earnings dynamics of men and women, that have been largely neglected in previous studies of earnings dynamics. The inequality among men is dominated by the permanent component. Conversely, permanent and transitory components are of comparable magnitudes to women. As a corollary, men experience more stable income paths but display larger permanent earnings differences. Women, on the other hand, face more unstable earnings profiles but show smaller permanent differences in earnings.

JEL codes: J31, J62

Keywords: Earnings distribution, earnings dynamics, permanent inequality, transitory inequality, variance decomposition

University of Helsinki and HECER.
Tel: +358-50-5323518, Fax: +358-9-19128736
Email: otto.kassi@helsinki.fi.

1 Introduction

Growing earnings inequality has been a common phenomenon to most of the developed countries since the 1970s and the need to understand this phenomenon has spurred a great deal of research.

Traditional studies of earnings inequality in Finland (e.g. Eriksson and Jäntti, 1997), as well as in other countries, have concentrated on measuring cross-sectional earnings inequality and its annual changes. However, concentrating on cross-sectional inequality hides an important element of economic inequality, namely the level of mobility of individuals within the earnings distribution.

More recent studies on earnings dynamics stress the importance of decomposing earnings inequality into its permanent and transitory parts. These two components have a different impact on long-term income differences and consequently have different welfare implications. If the rise in annual income inequality is driven by the transitory component, it suggests that earnings have become more volatile. This, in turn, may lead to a decrease in welfare, if individuals are unable to completely smoothen out income fluctuations. This might happen if earnings shocks are either very large or very persistent. On the other hand, if the rise in annual income inequality is due to fixed worker attributes, it implies that there is also increased inequality in career earnings. If the annual income inequality is driven by the transitory component, we should observe more year-to-year mobility within income distribution. This would lead to an increase of inequality in the short term; however it would even out in the long run. If the permanent component dominates the transitory, low earnings are a permanent rather than isolated experience.

Examples of factors contributing to the permanent component of earnings include changes in returns to education or skills, on-the-job training, or other factors that are relatively fixed from the point of view of an individual worker.[1]

In this paper, I decompose the annual variance of earnings into permanent and transitory components and study their evolution over time by fitting an error component model to observed second moments of earnings processes using Finnish data. My data are based on filed tax reports, so measurement errors due to misreporting are arguably substantially smaller than in survey-based approaches.

The vast majority of existing studies on earnings dynamics concentrate solely on males, thereby making the implicit assumption that earnings inequality between male workers is a good measure for overall earnings inequality.[2] The main contribution of this paper to the existing literature is that I present the decomposition of earnings separately for men and women. My approach echoes the observations of Korkeamäki and Kyyrä (2006), who, using Finnish data, found substantial differences in the educational background between men and women and also observed that men occupations and firms tended to be segregated into those that were dominated by males and those by females. Consequently, a picture of earnings inequality solely based on males might

[1] It should be stressed that income volatility may or may not be equivalent to economic risk. As discussed in Blundell et al (2008), earnings volatility does not necessarily translate into changes in welfare. Whether changes in earnings volatility have welfare implications depends on whether changes are anticipated and whether individuals are able to insure themselves against instability of earnings.

[2] A notable exception is Ziliak et al (2011), who report measures of permanent and transitory earnings inequality separately for men and women and for different educational groups, but do not limit their study to employed people.

be misleading. To get comparable figures for men and women, I limit my sample to working males and females and compare their earnings dynamics. Finally, my earnings data span the years 1988-2007, allowing me to study relatively recent developments in earnings dynamics.

My paper is heavily influenced by a series of articles that study earnings dynamics in other countries. Pioneering studies in this field include Gottschalk and Moffitt (1994), Moffitt and Gottschalk (2002), Baker (1997), and Haider (2001), all of which study earnings dynamics in the U.S. Following in their footsteps, Baker and Solon (2003) and Dickens (2000) present similar decompositions for Canada and the U.K., respectively. Due their access to a larger data set, they are able to fit more general models than the ones based on U.S. data. More recent papers using European registry based data fit variants of Baker and Solon (2003) and Dickens (2000). These include Gustavsson (2008), who studies Swedish panel data from 1960 to 1990, Ramos (2003) who studies British earnings data from the 1990s and Cappellari (2004), who studies Italian earnings data from the 1970s to 1990s. Even though the exact model specifications and time periods under consideration vary from country to country, the general finding is that there are significant differences between countries in terms of earnings dynamics. It is not clear whether the differences can be attributed to prevailing institutions or differences in the data. This creates a need to replicate the analysis using data from a new country. This paper is a scientific replication study (using the terminology of Hamermesh 2007): it applies a rather well-established model to a new data.

To give a preview of the results, it transpires that increasing earnings inequality is driven by both the permanent and the transitory components; however their contribution is different for men and women. For men, permanent inequality predominates over transitory inequality. For women, they are of a similar magnitude. In addition, permanent earnings differences vary substantially between cohorts. Male cohorts are more equal in terms of their permanent earnings . There has also been a trend of increasing earnings instability for both sexes during the observation period.

This paper is structured as follows: Section 2 describes the data and the sample selection criteria applied. Section 3 introduces the model of earnings dynamics and outlines the estimation method. Section 4 provides the results and subsequent discussion. Section 5 contrasts the findings to previous studies. Section 6 offers conclusions.

2 Data and sample construction

The data consists of a panel of a one-third random sample of Finnish census. It covers the years 1988-2007.

The measure of earnings used in this paper is annual gross income. Earnings are calculated from individual tax files. To ensure comparability, all earnings are deflated to EUR 2007 using the Consumer Price Index. By definition, annual earnings are given by hourly wage multiplied by hours worked. Therefore, the observed earnings inequality reflects two dimensions of inequality, inequality in wages and inequality in hours worked. Consequently, the variance of annual earnings is higher than the variance of hourly wages unless the covariance of wages and hours worked is negative and large (Abowd and Card, 1989).

My measure of earnings inequality is the variance of log annual labor earnings. Using the variance of log earnings as a dependent variable is a standard approach in papers studying earnings dynamics because the mathematical properties of variance are well established. In addition, the correlation between the variance of log earnings and other widely used inequality measures is very high. The downside of this choice is that it is not measure-free. Thus, the choice of currency unit and base year affects the measure of total earnings inequality. Nonetheless, the measure only affects the *level* of inequality, not the *changes*. Moreover, the decomposition into permanent and transitory components is unaffected by the measure.

Registry data has some advantages over survey data. Since the earnings information is collected by the authorities as a part of an administrative process, non-response and incorrect answers can be ruled out, which results in extremely reliable data on earnings.[3] Attrition from the data can occur only by migration or death. In addition, the definition of taxable labor earnings has remained unchanged for the period of observation.

Naturally, concentrating solely on labor earnings hides some of the income differences prevalent in the society. However, I have chosen this approach because supplementing the data by including capital income is not feasible due to limited available data. Moreover, including income transfers and paid taxes would introduce problems, because changes in tax laws and social security eligibility rules would severely limit the length of the panel. Another reason to prefer the measure of income chosen in this paper is that it is broadly equivalent to other papers published on the topic, thus facilitating international comparisons.

Another minor caveat in the data for the purposes of this paper is that earnings of over 200,000 Euros are top-coded due to statistical secrecy laws. This group is small (between .01 % and .05% of yearly observations), so their effect on the results is arguably small.

2.1 Sample selection criteria

The sample selection criteria were adapted from Haider (2001). They were chosen to ensure that the earnings dynamics of individuals in work are not confounded by people switching between work and non-work.

The target group in my sample is working males and females of prime age age between 26 and 60, who are observed for at least six years. I assume that by the age of 26, most people have completed their highest degrees.

I only include person-year observations if the main type of activity of a person is "working." In other words, I exclude students, the unemployed, the retired, and other people outside the workforce. I limit my attention to people who are working because my interest is in the earnings dynamics of people who are above the extensive margin. I also exclude working people with zero yearly earnings, as these observations are likely to have been misclassified.

After applying the sample selection criteria, I am left with a "revolving unbalanced panel" (following the terminology of Haider, 2001). The panel is unbalanced because all the cohorts are not observed for all the years. The length of the panel varies between 6 and 20 years, depending on the cohort. Using an unbalanced panel breaks the collinearity between year, age, and cohort effects, making it possible to identify them separately. Since people are only included if they fulfill the selection criteria, they may

[3] Gottschalk and Huynh (2010) show that earnings inequality decompositions based on U.S. survey data most likely overstate total inequality due to non-classical measurement errors.

enter and exit the panel. This feature makes the panel revolving. Applying a revolving, unbalanced panel mitigates problems related to compositional changes in the workforce due to the business cycle. If workers with unstable earnings only enter the workforce during an economic boom, they are only included in the data for those years, for which other selection criteria are fulfilled.

Since the individuals with very volatile earnings are also more likely to permanently exit the panel, the approach chosen here introduces a potential selectivity bias to the estimates. Correcting for attrition is not feasible because the data lack instruments for selection. Still, the approach chosen here is less restrictive than analyses based on fully balanced panels. In addition, only including people with no breaks in their earnings histories would probably overstate the contribution of the permanent earnings component.

Previous papers studying the covariance structure of earnings concentrate solely on males. The underlying assumption behind this is that the labor force participation of men is more or less constant, whereas female labor force participation is jointly determined with family decisions (e.g. fertility), which may bias the results. Using a revolving balanced panel partially mitigates this problem, because only observations from working years are included. Therefore, transitions into and out of the workforce do not contribute to the empirical estimation. Nonetheless, it might be be the case, that the working hours of females vary more than those of males, which may be reflected in female earnings variances. In addition, it is well established, both theoretically and empirically (see, e.g., Eckstein and Wolpin, 1989; Euwals et al, 2011), that a large negative earnings shock may promote female fertility decisions. Fertility decisions might then lower female wages due to their effect on work experience of women. This mechanism introduces a specific kind of selectivity issue: women with high earnings shocks may voluntarily drop out of the workforce and concentrate on home production.[4] Notwithstanding these caveats, the data should be representative of those women who are well attached to the labor force. Furthermore, the labor force participation rate of Finnish women is very high (Pissarides et al, 2003), which means that the endogenous participation of women is less of a problem than in some other countries.

A revolving balanced panel structure ensures that the measure of earnings inequality in this paper reflects the true earnings inequality of the population with good attachment to the labor market. Even though sample selection criteria somewhat differ from other studies, due to different structure of the data used, they are consistent within the observation period, thus enabling comparisons between years. Comparisons between countries, on the other hand, might be more questionable.

I categorize people into two year birth cohorts and follow each cohort through time. Studies based on a smaller data have been forced to pool all cohorts together due to small sample sizes. This naturally hides some of the heterogeneity of earnings dynamics between cohorts. The total size of the sample used in the analysis is given in Table 1.

2.2 Descriptive statistics of the covariance structure

In Figure 1, I plot the observed earnings variance for workers selected by the selection criteria given above. For both sexes, the variance decreases between the years 1988-1991

[4] It should be noted, that a similar mechanism might be present for male workers too: a large negative earnings shock may also induce men to drop out of the workforce.

and thereafter rises until reaching its peak around 1994. After 1994 earnings inequality falls somewhat but remains high until the end of the sample period. The variances plotted in Figure 1 are somewhat higher than those observed in most other similar studies. This might be because I cannot discriminate between full-time and part-time workers. Moreover, in some studies based on income tax reports, earnings are censored from below, because income below the tax limit is not observed. This is not the case in this paper.

To grasp the essential features of earnings dynamics, it is useful to inspect the autocorrelation profiles of earnings by year and cohort. I have calculated yearly variance and autocovariances between years for people who are observed in both years. For cohorts who are observed for the full twenty years this adds up to 210 unique covariance elements ($21 \times 20/2$) and less for the other cohorts. In total, the unique elements of covariance matrices add up to 3,066 covariance elements.

Figure 2 presents the yearly variances and covariances between annual earnings for selected cohorts of men and women. Figure 2 demonstrates that there are substantial differences in the variances and autocovariances of male and female earnings. This suggests that there are considerable differences in the earnings dynamics of men and women, making it reasonable to estimate separate models for the two sexes. In addition, a comparison of years reveals strong year effects. These are especially apparent during the recession of the early 1990s. The difference between variance and the first autocovariance is relatively large. In addition, autocovariances remain positive even at long lags, indicating that there are considerable permanent earnings differences. Finally, the variance and autocovariance values are larger for the oldest cohort, even at longer lags, which suggests the presence of cohort effects in the permanent component of earnings.

An alternative way to study cohort covariances is to keep the year fixed and plot covariances by age. This is done for three selected years in Figure 3. Comparing years reveals that income variances and covariances have risen over time for both men and for women, which indicates that earnings inequality has increased during the panel time and that at least part of this rise is due to a rise in permanent earnings differences. The variances are higher for young women than for young men, but as people grow older, the higher growth in variances of male earnings causes men to overtake women in terms of earnings inequality. The difference between the variance and autocovariances of earnings is at its largest for young women, indicating that high earnings inequality among young women is driven by transitory differences. For men, the difference between the variance and covariances remains almost constant, regardless of age.

To summarize, in addition to being able to disentangle permanent and transitory income differences, the preferred model for earnings inequality should allow reflect both cohort and year effects. The model should also allow variances of permanent and transitory components to change as people age.

3 Model and estimation

In this section, I introduce the econometric model and the estimation method. I estimate an error-components model with permanent and transitory components. The model allows individuals to permanently differ in their mean earnings as well as the earnings growth rate. The transitory component is modeled as an AR(1) process. As a result, even transitory shocks are allowed to exhibit persistence and can consequently take more than one year to flatten out.

3.1 Econometric model

Let Y_{ibt} denote log earnings in year t of person i born in year b. Individual earnings can be expressed as deviations from means, or

$$Y_{ibt} = \mu_{bt} + y_{ibt}.$$

Since my interest lies in the second moments of the distribution of Y_{ibt}, it suffices to write a model for de-meaned wages y_{ibt}. Expressing μ_{bt} as cohort-age means captures the average year, age, and cohort effects in a more flexible fashion than using regression models with cohort-specific polynomials. The simplest possible model for y_{ibt} is

$$y_{ibt} = p_t \alpha_{ibt} + \lambda_t \varepsilon_{ibt} \tag{1}$$

where the two terms are assumed to be orthogonal to each other. Equation (1) can be seen as a Mincerian earnings equation of relative earnings, where α_{ibt} stands for the observed characteristics of individuals and ε_{ibt} is the error term. p_t and λ_t are year-specific factor loadings. Applying the variance operator to both sides yields

$$\mathsf{Var}\,(y_{ibt}) = p_t^2 \sigma_\alpha^2 + \lambda_t^2 \sigma_\varepsilon^2. \tag{2}$$

Equation (2) gives the basic intuition of the decomposition. $p_t \sigma_\alpha^2$ denotes the variance of the permanent component of earnings, and $\lambda_t \sigma_\varepsilon^2$ denotes the variance of the transitory component. An increase in either component increases the dispersion of earnings, but an increase in $\lambda_t \sigma_\varepsilon^2$ also implies that churning within the earnings distribution increases.

Even though Equation (2) is intuitive, it may be too restrictive for two reasons. First, the variance of transitory shocks may exhibit age-related heteroskedasticity because workers at the start of their careers may have more unstable earnings.[5] In addition, different cohorts may have different skills or other idiosyncratic features that affect the variability of their earnings. To incorporate these features, the following generalization of Equation (1) is used

$$y_{ibt} = q_c p_t u_{ibt} + \varepsilon_{ibt}, \tag{3}$$

where

$$u_{ibt} = \alpha_i + \beta_i x, \tag{4}$$

$$\varepsilon_{ibt} = \rho \varepsilon_{ibt-1} + \lambda_t \nu_{ibt}. \tag{5}$$

The terms α_i, β_i and ν_i are random variables with distributions:

$$\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} \sim \left(\begin{bmatrix} \bar{\alpha} \\ \bar{\beta} \end{bmatrix}, \begin{bmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ & \sigma_\beta^2 \end{bmatrix} \right) \tag{6}$$

[5] Meghir and Pistaferri (2004) report significant heteroskedasticity in earnings instability using U.S. data, albeit using a different model. Baker and Solon (2003) and Gustavsson (2008) reach similar conclusions using Canadian and Swedish data, respectively.

and

$$\nu_{ibt} \sim \left(\bar{\nu}, \gamma_0 + \gamma_1 x + \gamma_2 x^2\right). \tag{7}$$

x is defined as the potential experience of each cohort in year t, i.e., $x = t - b - 26$.

In Equation (4) u_{ibt} is a random growth term. It describes the permanent component of earnings. σ_α^2 reflects the variance of the earnings profiles of individuals at the age of 26, and the variance in σ_β^2 reflects the deviation of the individual-specific growth rate from the average growth rate of each cohort (the average growth rate is captured in μ_{bt}). p_t and q_c are year and cohort factor loadings, respectively.

The transitory component of earnings in (5) is given by a mean-reverting AR(1) process. λ_t are year-specific factor loadings on the innovation ν_{ibt}. This specification assumes that an earnings shock takes more than one year to flatten out and that earnings shocks accumulate over time. In addition, Equation (7) allows transitory variance to be a quadratic function of age. The transitory and permanent components of earnings are assumed to be orthogonal to one another. To make identification possible, I have normalized p_{1988} and λ_{1989} and $q_{1951-1952}$ to 1.

Equations (3)–(7) generate non-stationarity to the variances of earnings processes through time-varying factor loadings for the permanent and the transitory components, p_t and λ_t. Another source of non-stationarity is the polynomial form of transitory shock variance. The intuition remains the same: a rise in p_t or q_c increases permanent income differences, whereas a rise in λ_t increases the shuffling of workers.

In line with the model specifications in Baker and Solon (2003) and Gustavsson (2008), polynomial form of $var(\nu_{ibt})$ recognizes that the earnings instability may vary between individuals because they are at different stages of their careers. The yearly factor loadings of the permanent and transitory components also give insights into the forces driving changes in income distribution.

The justification for the formulation in Equation (4) is both theoretical and empirical. For example, it has been successfully applied in Haider (2001), Ramos (2003), and Cappellari (2004), who demonstrate that in addition to allowing heterogeneity in mean earnings, the slope of earnings and the covariance of the two are important for capturing the dynamics of earnings. In most previous studies, the covariance term $\sigma_{\alpha\beta}$ is found to be negative. This is consistent with the on-the-job training hypothesis (see, e.g., Lillard and Weiss 1979; Hause 1980; and Baker 1997), which states that individuals may accept lower earnings at the beginning of their career, since they anticipate that their earnings will rise at a high enough rate and for a long enough time to compensate for low earnings at the beginning of their career. On the other hand, if $\sigma_{\alpha\beta}$ is found to be positive, it is consistent with the schooling-matching hypothesis, in which better skilled workers are endowed with more education, which raises their initial earnings and causes them to enjoy faster earnings growth as the quality of the match is revealed to their employers (Cappellari, 2004).

In addition to the specification in Equations (3)–(7), usually known as "random growth specification", I have also experimented with other specifications, in particular, with the so called "random walk specification" (e.g., (Gustavsson, 2008)), which is another widely used specification. This model is given by $u_{ibt} = u_{ibt-1} + \xi_{it}$, where ξ_{it} is a white noise process. The main difference between the two formulations is that the random growth specification allows the correlation of the intercept and slope terms to be nonzero, whereas the random walk specification does not. In this sense, the random growth specification nests the random walk specification. Trials with the random growth specification always resulted in a statistically significant estimate for $\sigma_{\alpha\beta}$. I interpret this as a sign that the random walk model is inconsistent with the observed covariance structure of earnings.

Random walk and random growth specifications have different implications in terms of the age-derivative of cross-sectional variances. Under the random walk specification, the variance of earnings increases linearly with age, whereas under the random growth specification, the growth of permanent earnings inequality is either convex or concave, depending on the sign of $\sigma_{\alpha\beta}$ (Guvenen, 2009).[6]

3.2 Estimation

Direct estimation of a model based on equations (3)-(7) is inefficient, because it means estimating α_i and β_i for each individual with only a small number of observations. Since I am interested in the second moments of earnings distribution, I therefore estimate them directly. To accomplish this, I write down the variance of earnings in year t for cohort b implied by (3):

$$\begin{aligned}\mathsf{Var}\left(y_{ibt}\right) = q_c^2 p_t^2 &\left[\sigma_\alpha^2 + x^2\sigma_\beta^2 + 2x\sigma_{\alpha\beta}\right] + \\ &\rho^2 var\left(\varepsilon_{ibt-1}\right) + \lambda_t^2 var\left(\nu_{ibt}\right). \end{aligned} \tag{8}$$

Respectively, a general covariance element between year t earnings and year $t-h$ ($h>0$) earnings is given by[7]

$$\begin{aligned}\mathsf{Cov}\left(y_{ibt}, y_{ibt-h}\right) = q_c^2 p_t p_{t-h} &\left[\sigma_\alpha^2 + x(x-h)\sigma_\beta^2 + (2x-h)\sigma_{\alpha\beta}\right] \\ &+\rho^h var\left(\varepsilon_{ibt-h}\right). \end{aligned} \tag{9}$$

The term $var\left(\varepsilon_{ibt}\right)$ is calculated by backtracking the recursion in Equation (5) to the first sample year of each cohort. Since the earnings time series are relatively short, consequent covariances depend on the variance of the initial shock. This makes the standard time series analysis assumption of zero initial conditions problematic. I follow the suggestion of MaCurdy (1982; 2007) and treat the variances of initial shocks as extra parameters to be estimated. This parameter also takes into account transitory earnings differences accumulated before the start of the sample.[8]

Since the panel is revolving, an individual can only contribute to an element $(t, t-h)$ in the covariance matrix, if he or she is observed in years t and $t-h$. The sample covariances are thus calculated as the earnings covariance of people who are observed in

---

[6] Since random walk and random growth specifications do not necessarily rule each other out, some researchers (e.g. Baker and Solon, 2003 and Ramos, 2003) incorporate both into the same model. In my data, this specification either does not converge or results in negative variance estimates. Furthermore, the interpretation of these nested specifications is far from clear.

[7] Identification of earnings instability is made possible only by the off-diagonal elements of the covariance matrix. The underlying intuition is that a high correlation between earnings at t and earnings at $t-h$ implies a low instability of earnings.

[8] Initial variance parameters have a different interpretation depending on whether earnings trajectories start from the age 26 or at a later age. For a cohort who has been 26 years of age before 1988, the initial variance is a measure of the transitory variance accumulated before 1988, whereas for a cohort who is observed for the first time after 1988, the initial variance is a measure of labor market conditions at the time of labor market entry.

both years. Consequently, people who have a higher attachment to the labor market contribute more to the empirical covariance matrices, which leads to a sample selection problem that cannot be completely overcome by unbalanced revolving panel construction. This is a common caveat in papers of this genre.

The estimation boils down to minimizing the distance between the scohort earnings covariances implied by the model and the empirical autocovariances calculated from the data. I stack each unique covariance matrix element into vector $\mathbf{C}$. The estimation is done by GMM, i.e., by minimizing the distance between observed autocovariances $\mathbf{C}$ and those implied by the model $\mathbf{F}(\theta)$, where θ is a vector of 87 parameters to be estimated. In practice, I minimize the standard GMM criterion function

$$H = [\mathbf{C} - \mathbf{F}(\theta)]' \, \mathbf{W} \, [\mathbf{C} - \mathbf{F}(\theta)] \,. \tag{10}$$

Altonji and Segal (1996) demonstrate that using the asymptotically optimal GMM weighting matrix, i.e., choosing $\mathbf{W} = \left[\mathbf{F}(\theta)' \, \mathbf{F}(\theta)\right]^{-1}$, can lead to a very large finite-sample bias. This is due to $\left[\mathbf{F}(\theta)' \, \mathbf{F}(\theta)\right]$ being very close to singular. In line with the bulk of the earnings dynamics literature, I have chosen the identity matrix as the weighting matrix. This approach is called the Equally Weighted Minimum Distance estimation (Chamberlain, 1984). Using the identity matrix as the weighting matrix gives consistent but generally inefficient estimates.

The asymptotic standard errors of vector θ are given by the standard GMM covariance matrix based on the fourth moments of the data. That is

$$\mathsf{Var}(\theta) = \left(\mathbf{D}'\mathbf{D}\right)^{-1} \mathbf{D}'\boldsymbol{\Omega}\mathbf{D} \left(\mathbf{D}'\mathbf{D}\right)^{-1},$$

where $\mathbf{D} = \dfrac{\partial \mathbf{F}(\theta)}{\partial \theta'}$ and $\Omega = [\mathbf{C} - \mathbf{F}(\theta)]' \, \mathbf{Q} \, [\mathbf{C} - \mathbf{F}(\theta)]$ are evaluated at the solution $\theta = \hat{\theta}$. $\mathbf{Q}$ is a block diagonal matrix of ones. Including $\mathbf{Q}$ in the matrix product effectively sets the covariances between cohorts at zero.

4 Estimation results

4.1 Parameter estimates

Figure 4 decomposes total earnings inequality into its permanent and transitory components. The decomposition is based on equation (8). The term $p_t^2 \left[\sigma_\alpha^2 + x^2\sigma_\beta^2 + 2x\sigma_{\alpha\beta}\right]$ accounts for the permanent component of earnings and term $\rho^2 var\left(\varepsilon_{ibt-1}\right) + \lambda_t^2 var\left(\nu_{ibt}\right)$ accounts for the transitory component.

The contribution of permanent earnings inequality to total inequality is larger for men than for women in almost all years. This implies that permanent inequality among men is larger than among women. The contribution of the permanent component has remained roughly similar throughout the sample, with the exception of the recession years 1991 and 1992. Even though the magnitude of the two components differ between sexes, the dynamics of the two components were roughly similar for both sexes for the entire sample period.

The transitory components of the earnings inequality of men and women are highly correlated, although the transitory inequality is higher for women. Tables 2 and 3 present the estimates of the parameters of the permanent and transitory components. I will discuss both parameters in turn.

An initial glance at Tables 2 and 3 shows that most parameter estimates have very small standard errors. That is, they are accurately estimated, in spite of the model being flexibly parameterized and including year, cohort, and experience effects. Table 2 reports the parameters of permanent earnings differences. The level term σ_α^2 is statistically significantly larger for men than it is for women. The slope term σ_β^2 and the correlation term $\sigma_{\alpha\beta}$ are of similar magnitude for both sexes. Moreover, the estimated correlation between the intercept and slope terms is positive. This means that people who have higher initial earnings also have larger earnings growth. As a result, the permanent part of earnings distribution becomes increasingly unequal over the life cycle.

For men, the yearly loadings on the permanent earnings component are almost constant, except for the two deepest recession years. For women, there is a downward trend in yearly loadings, indicating that permanent earnings differences have decreased during the end of the 1990s and early 2000s. Changes in the permanent component yearly loading can be interpreted as the prices of the fixed skills of individuals, keeping cohort effects constant.

The deep Finnish recession experienced by Finland in the early 1990s is visible as a drop in the permanent earnings inequality component.[9] The explanation for this drop is that people with lowest wages dropped out of the workforce, which had the effect of decreasing earnings inequality. However, the effect of the recession is much less clear in the time series of transitory shocks.

Next, I will turn to estimates for cohort loadings on the permanent component q_c. The most intuitive interpretation for the cohort loadings of permanent component is that they are a measure of the dispersion of the skills within a cohort. An alternative interpretation for q_c is that they reflect very persistent shocks that affect cohorts differently even if the skill dispersion of cohorts does not change. An example of such shocks is the long-term wage scarring effect of graduating in a recession (see, e.g., Kwon et al 2010, and Oreopoulos and von Wachter 2008).

Coinciding with the decrease in yearly factor loadings for women is the increase in cohort loadings for the permanent component of earnings. This implies that earnings inequality for all women has decreased since the 1990s, but at the same time, younger cohorts are more unequal than older cohorts. In other words, as younger cohorts have become more skilled, within cohort inequality has increased but at the same time the inequality between cohorts has decreased. For men, however, the inequality time-trend is exactly the opposite: the younger cohorts are more equal than the older ones.

Finally, I will turn to the estimates for the transitory component reported in the bottom four rows of Table 3. The parameter of the autocorrelation, which is estimated at 0.22 for men and women. The bottom three lines in Table 3 give the parameters of age-heteroskedasticity. The age profile of variance of the transitory shocks, visualized in Figure 5, is strikingly different between men and women. For men, γ_1 and γ_2 do not differ from zero at conventional risk levels, which implies that there is no age-related heteroskedasticity in the variance of transitory shocks. For women, on the other hand, the variance of transitory shocks is decreasing and convex. For women under 30, transitory shock variance is over double that of men. Furthermore, regardless of the cohort, initial earnings shocks are considerably higher for women than for men.

[9] Finland experienced the deepest economic recession experienced in any industrialized country since the 1930s. For details, see, e.g., Gorodnichenko et al (2009) and Honkapohja and Koskela (1999).

This observation is roughly consistent with Lundberg and Rose (2000), who find that motherhood decreases the labor supply of married women who are attached to the labor market but not their wages.

Yearly loadings on the transitory component of earnings λ_t also exhibit different trends for men and women.[10] For men, they in 1994, whereafter, they decline somewhat but still remain above one until 2007. For women, there is an almost constant rise from 1988 to 2007.

Adding cohort loadings to the transitory component always resulted in convergence problems. This suggests that cohort effects on the transitory component overparameterize the model. Therefore, earnings instability seems to be symmetric for all cohorts when initial conditions have been accounted for. Earnings instability seems to be more related to labor market conditions prevailing in society than to differences in the human capital within cohorts. Nonetheless, for both sexes, the contribution of rising earnings instability to inequality is substantial.

4.2 Decomposition analysis: cohorts and years

The parameter estimates only give a partial description of the evolution of earnings dynamics. As discussed in the previous subsection, there is substantial variation between cohorts and years. To get further insight into these differences, this subsection introduces counterfactual analyses, which are obtained by eliminating sets of parameters in turn.

Figure 6 plots the contributions of the cohort and the year effects on permanent inequality. For men, setting the year effects to 1 eliminates most of the permanent differences. This is consistent with the notion that permanent earnings differences of males are driven by the changes in returns to skill rather than differences in the skill composition of cohorts. However, the same explanation does not apply for females, as eliminating the year effects in permanent component actually increases female earnings inequality. This suggests that permanent female inequality is driven by within cohort inequality rather than year-to-year changes.

Turning to transitory earnings differences, we see similar patterns for men and women. Eliminating the year loadings flattens most of the transitory shocks. This unequivocally suggests that earnings have become more unstable for both sexes. The slight downward trend in the transitory inequality of females is explained by the age-gradient of the transitory shock variance: older people face smaller transitory shocks.

The underlying assumption in the preceding discussion is that the model is correctly specified and all of the parameters are strongly identified. The following section discusses the evidence in favor of strong identification.

4.3 Sensitivity of results to model specification

Generally, weak identification can arise if the moment condition is small, but not zero, at a range of values differing from the true parameter value θ_0. Stock and Wright

[10] So that the initial variance parameter are only identified by the initial variance in a cohort's first sample year, λ_{1988} is left unrestricted and λ_{1989} is normalized to 1. Without this restriction, the yearly loadings on the transitory component and the initial variances could not be jointly identified.

(2000) show that the asymptotic theory devised for identified models is invalid for weakly identified models. As a result, the parameter estimates of weakly identified models are inconsistent, and the calculated covariance matrix does not converge to the true covariance matrix, which results in invalid estimates for standard errors. Furthermore, even if most parameters are strongly identified, their asymptotic standard errors might be invalid in the presence of some weakly identified parameters. In the context of this paper, weak identification may arise if ρ is close to 1. If $\rho \approx 1$, the transitory component is "very close to being permanent". This causes problems in identification, since both the transitory and permanent components reflect relatively permanent earnings inequality making it difficult to distinguish them from one another, especially if the panel length is short.

Doris et al (2012) gives Monte Carlo evidence on the ranges of parameter values that lead to biased estimates. According to Tables 2a and 3 in their paper, a model estimated using eight panel years of observations and $\rho = .8$ is sufficient to give unbiased results. Since my panel length is well over that for most cohorts (median panel length in my data is 17 years) and my estimates of ρ are well below .8, I am confident about the strong identification of the models estimated. In addition, linear time trends in any factor loading might also cause problems in identification. Such trends are not present in my data.

I have not applied Newey's (1985) specification test to assess the goodness of fit, because a general finding in the earnings dynamics literature is that the null hypothesis of a correctly specified model is almost always rejected. According to Baker and Solon (2003), these tests have inflated sizes when the amount of overidentifying restrictions is as large as in this case (3,066 moment conditions used to identify 87 parameters).

Another possible caveat, according to the warnings given in Baker and Solon (2003) and Shin and Solon (2011), is that an arbitrary change in a parametric model may lead to different conclusions. I have experimented with alternative specifications (given in Table 4) and do not find cause for concern. First, as is evident from the difference in the mean squared errors (MSE's) between the random growth and random walk models and the statistically significant estimate of $\sigma_{\alpha\beta}$, the data clearly reject the simpler random walk specification in favor of the random growth specification.

Second, the model with an ARMA(1,1) specification for the transitory component gives qualitatively similar results to the AR(1) specification. The main difference in these two specifications is that the inclusion of the MA(1) parameter reduces the absolute value of the persistence parameter ρ. This difference is statistically significant for women but not for men. Inclusion of the MA(1) parameter does not have a jointly statistically significant impact for other parameters other than ρ.

5 Comparison to other studies

To better grasp which of the differences between this paper and papers using data from other countries are due to differences in econometric modeling choices or data and which are due to prevailing institutional differences, this section contrasts the central findings of this with previous studies. Two findings in this paper are qualitatively different from most previous papers: the sign of the correlation between earnings growth and intercept components, $\sigma_{\alpha\beta}$, and the autocorrelation coefficient of the transitory earnings component ρ.

My estimate for $\sigma_{\alpha\beta}$ is positive for both sexes, which implies that the distribution of earnings widens as people age. Cappellari (2004) reports a similar finding for Italy, but other studies I am aware of (Haider, 2001; Baker and Solon, 2003; Baker, 1997; Gustavsson, 2008; Ramos, 2003) report a negative parameter estimate. It is difficult to judge whether these differences are more due to data construction or institutional differences. Nonetheless, it should be noted that the studies that find $\sigma_{\alpha\beta} < 0$ use various earnings measures as well as data sources.

Since I do not observe the hours worked by an individual, I am not able to discriminate between full-time and part-time workers. This inevitably affects some of the results. My estimates for the contribution of the transitory component to total inequality are much larger than those obtained in studies concentrating only on full time workers (these include Baker and Solon, 2003; Haider, 2001; Gustavsson, 2008). Moreover, the persistence of transitory shocks is found to be considerably lower. Baker and Solon (2003), Dickens (2000), and Haider (2001) report estimated autocorrelation values in the range of 0.6 and 0.95, however, they concentrate on full-time working males. In contrast, Ramos (2003) does not discriminate between full-time and part-time workers. He finds that the transitory component may account for up to 80 percent of yearly earnings variance. Ramos also finds $\hat{\rho}$ estimates in the range of 0.29 and 0.41, which are considerably closer to my estimates.

Another partial explanation to the low estimated persistence of transitory shocks is that, generally speaking, random walk specification results in higher persistence compared to random growth specifications. Indeed, Guvenen (2009) shows analytically that if a random growth model is misspecified as a random walk model, the persistence parameter ρ will be biased upwards.

Making comparisons between other countries is also somewhat suspect, because separating prevailing differences in labor market conditions from differences in the data is far from straightforward. The most comparable study to the current one in terms of data is Ramos (2003), who studies male earnings inequality in the U.K. between 1991 and 1999. Most notable difference between the results in Ramos (2003) and those in this that Ramos finds that older workers face very unstable earnings compared to younger workers. For example, according to his estimates, the transitory component constitutes over 80 per cent of the total income variance for a cohort over 50 years old at the end of the observation period. This is considerably higher than my findings; and it is most likely attributable to institutional differences (e.g., a higher proportion of part-time workers over 50 in Great Britain compared to Finland).

6 Summary and conclusions

Previous research has shown that earnings inequality has risen in Finland during the last two decades. This paper decomposed the yearly Finnish log earnings variance of the working population into its permanent and transitory components. The analysis is done separately for men and women. The econometric analysis is based on the second moments of log-earnings distribution using the Equally Weighted Minimum Distance method of Chamberlain (1984).

I find that the increase in earnings inequality among men and women is driven by both the permanent and transitory components, but the contributions of these components are of different magnitude. The permanent component of earnings inequality is larger for men than for women. As a corollary, men enjoy more stable income paths but with larger permanent earnings differences. Women, on the other hand, experience more unstable earnings processes but have smaller permanent differences in earnings.

The age-derivative of the permanent earnings inequality of men and women is similar, indicating that the relative differences in permanent earnings stay similar throughout the careers of men and women. The correlation between initial earnings inequality and the growth in earnings inequality is found to be positive for both sexes, implying a divergence of earnings profiles and increasing permanent earnings differences toward the end of individuals' working career. Compared to findings, in other countries the persistence of transitory earnings shocks is found to be relatively small. Moreover, the contribution of transitory shocks to inequality has risen considerably for both sexes. This strongly suggests that earnings have become more unstable during the last 20 years.

Finding ultimate causes for the changes in persistent and transitory inequality is beyond the scope of this paper, but some tentative explanations can be offered. For both sexes, we see that year loadings on the permanent inequality drive a lot of the earnings differences. This might be due to yearly changes in labor demand. On the other hand, the larger cohort effects of younger cohorts on permanent female earnings inequality suggest that younger working women face higher permanent earnings inequality than older women. It seems plausible that this is due to the high labor force participation of young women rather than changing returns to skill, as there is no such trend for cohorts of men.

Finally, the lessons of this paper suggest that researchers applying estimates obtained from these types of models in their work may inadvertedly miss potentially important aspects of the earnings dynamics prevalent in society if they only concentrate on males.

Acknowledgements

I am thankful to the two anonymous referees, Bernd Fitzenberger (the editor), Gero Dolfus, Anssi Kohonen, Heikki Pursiainen, Tatu Westling and the seminar participants at the HECER Econometrics and Computational Economics seminar for various useful comments and to Merja Kiljunen and Marjo Pyy-Martikainen from Statistics Finland for helping me with the data. Financial support from the Kone Foundation and OP Pohjola Foundation is gratefully acknowledged.

References

Abowd JM, Card D (1989) On the covariance structure of earnings and hours changes. Econometrica 57(2):411–45, URL `http://ideas.repec.org/a/ecm/emetrp/v57y1989i2p411-45.html`

Altonji J, Segal L (1996) Small-sample bias in gmm estimation of covariance structures. Journal of Business and Economic Statistics 14(3):353–366

Baker M (1997) Growth-rate heterogeneity and the covariance structure of life-cycle earnings. Journal of Labor Economics 15(2):338–75, URL `http://ideas.repec.org/a/ucp/jlabec/v15y1997i2p338-75.html`

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

Blundell R, Pistaferri L, Preston I (2008) Consumption inequality and partial insurance. The American Economic Review 98(5):pp. 1887–1921, URL http://www.jstor.org/stable/29730156

Cappellari L (2004) The dynamics and inequality of italian men's earnings. Journal of Human Resources 39(2):475 – 499

Chamberlain G (1984) Panel data. In: Griliches Z, Intriligator MD (eds) Handbook of Econometrics, Handbook of Econometrics, vol 2, Elsevier, chap 22, pp 1247–1318, URL http://ideas.repec.org/h/eee/ecochp/2-22.html

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

Doris A, O'Neill D, Sweetman O (2012) Identification of the covariance structure of earnings using the gmm estimator. The Journal of Economic Inequality pp 1–30, DOI 10.1007/s10888-012-9216-5, URL http://dx.doi.org/10.1007/s10888-012-9216-5

Eckstein Z, Wolpin KI (1989) Dynamic labour force participation of married women and endogenous work experience. Review of Economic Studies 56(3):375–90, URL http://ideas.repec.org/a/bla/restud/v56y1989i3p375-90.html

Eriksson T, Jäntti M (1997) The distribution of earnings in finland 1971-1990. European Economic Review 41(9):1763 – 1779

Euwals R, Knoef M, Vuuren D (2011) The trend in female labour force participation: what can be expected for the future? Empirical Economics 40(3):729–753, URL http://ideas.repec.org/a/spr/empeco/v40y2011i3p729-753.html

Gorodnichenko Y, Mendoza EG, Tesar LL (2009) The finnish great depression: From russia with love. IZA Discussion Papers 4113, Institute for the Study of Labor (IZA), URL http://ideas.repec.org/p/iza/izadps/dp4113.html

Gottschalk P, Huynh M (2010) Are earnings inequality and mobility overstated? the impact of nonclassical measurement error. The Review of Economics and Statistics 92(2):302–315, URL http://ideas.repec.org/a/tpr/restat/v92y2010i2p302-315.html

Gottschalk P, Moffitt R (1994) The growth of earnings instability in the u.s. labor market. Brookings Papers on Economic Activity 25(1994-2):217–272

Gustavsson M (2008) A new picture of swedish earnings inequality: Persistent and transitory components, 1960-1990. Review of Income and Wealth 54(3):324–349, DOI 10.1111/j.1475-4991.2008.00278.x

Guvenen F (2009) An empirical investigation of labor income processes. Review of Economic Dynamics 12(1):58–79, URL http://ideas.repec.org/a/red/issued/06-15.html

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

Hamermesh DS (2007) Replication in economics. Working Paper 13026, National Bureau of Economic Research, URL http://www.nber.org/papers/w13026

Hause JC (1980) The fine structure of earnings and the on-the-job training hypothesis. Econometrica 48(4):pp. 1013–1029, URL http://www.jstor.org/stable/1912945

Honkapohja S, Koskela E (1999) The economic crisis of the 1990s in finland. Economic Policy 14(29):399–436, DOI 10.1111/1468-0327.00054, URL http://dx.doi.org/10.1111/1468-0327.00054

Korkeamäki O, Kyyrä T (2006) A gender wage gap decomposition for matched employer-employee data. Labour Economics 13(5):611–638, DOI DOI:10.1016/j.labeco.2005.02.001, URL http://www.sciencedirect.com/science/article/B6VFD-4G94HGV-1/2/b1853c12cc24380d0c0bb1c475a4ebeb

Kwon I, Milgrom EM, Hwang S (2010) Cohort effects in promotions and wages: Evidence from sweden and the united states. J Human Resources 45(3):772–808, URL http://jhr.uwpress.org/cgi/content/abstract/45/3/772, http://jhr.uwpress.org/cgi/reprint/45/3/772.pdf

Lillard LA, Weiss Y (1979) Components of variation in panel earnings data: American scientists 1960-70. Econometrica 47(2):pp. 437–454, URL http://www.jstor.org/stable/1914192

Lundberg S, Rose E (2000) Parenthood and the earnings of married men and women. Labour Economics 7(6):689 – 710, DOI DOI:10.1016/S0927-5371(00)00020-8, URL http://www.sciencedirect.com/science/article/pii/S0927537100000208

MaCurdy T (2007) Chapter 62 a practitioner's approach to estimating intertemporal relationships using longitudinal data: Lessons from applications in wage dynamics. Handbook of Econometrics, vol 6, Part 1, Elsevier, pp 4057 – 4167, DOI DOI:10.1016/S1573-4412(07)06062-X, URL http://www.sciencedirect.com/science/article/B7GX7-4R9GYJT-5/2/15c98f000e2780d9c077a6d609bbd225

MaCurdy TE (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, DOI DOI:10.1016/0304-4076(82)90096-3, URL http://www.sciencedirect.com/science/article/B6VC0-45PMHBS-6/2/eac62282167901d3ca7968ed7ac65ff9

Meghir C, Pistaferri L (2004) Income variance dynamics and heterogeneity. Econometrica 72(1):1–32, DOI 10.1111/j.1468-0262.2004.00476.x, URL http://dx.doi.org/10.1111/j.1468-0262.2004.00476.x

Moffitt RA, Gottschalk P (2002) Trends in the transitory variance of earnings in the united states. The Economic Journal 112(478):C68–C73

Newey WK (1985) Generalized method of moments specification testing. Journal of Econometrics 29(3):229 – 256, DOI DOI:10.1016/0304-4076(85)90154-X, URL http://www.sciencedirect.com/science/article/B6VC0-459J6SV-15/2/7c252f4694f56e162907c5400e93e4cd

Oreopoulos P, von Wachter T (2008) The short- and long-term career effects of graduating in a recession: Hysteresis and heterogeneity in the market for college graduates. IZA Discussion Papers 3578, Institute for the Study of Labor (IZA)

Pissarides C, Garibaldi P, Olivetti C, Petrongolo B, Wasmer E (2003) Women in the labour force: How well is europe doing? Open access publication, Sciences Po

Ramos X (2003) The covariance structure of earnings in great britain, 1991-1999. Economica 70(278):353–374, DOI 10.1111/1468-0335.00328, URL http://dx.doi.org/10.1111/1468-0335.00328

Shin D, Solon G (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, DOI DOI:10.1016/j.jpubeco.2011.02.007, URL http://www.sciencedirect.com/science/article/pii/S0047272711000338

Stock JH, Wright JH (2000) Gmm with weak identification. Econometrica 68(5):1055–1096, DOI 10.1111/1468-0262.00151, URL http://dx.doi.org/10.1111/1468-0262.00151

Ziliak JP, Hardy B, Bollinger C (2011) Earnings volatility in america: Evidence from matched cps. Labour Economics DOI DOI:10.1016/j.labeco.2011.06.015, URL http:

//www.sciencedirect.com/science/article/pii/S0927537111000765

Cohort	Years observed	Age in initial year	Sample size (men)	Sample size (women)
1933-1934	1988-1994	55	2 882	3 673
1935-1936	1988-1996	53	5 169	6 579
1937-1938	1988-1998	51	7 383	8 954
1939-1940	1988-2000	49	8 595	9 987
1941-1942	1988-2002	47	10 398	11 736
1943-1944	1988-2004	45	11 593	12 877
1945-1946	1988-2006	43	16 817	18 481
1947-1948	1988-2007	41	18 760	19 801
1949-1950	1988-2007	39	18 026	19 311
1951-1952	1988-2007	37	17 735	18 784
1953-1954	1988-2007	35	17 643	18 982
1955-1956	1988-2007	33	18 377	18 926
1957-1958	1988-2007	31	17 610	17 769
1959-1960	1988-2007	29	18 060	17 562
1961-1962	1988-2007	27	18 447	16 932
1963-1964	1989-2007	26	18 720	16 700
1965-1966	1991-2007	26	17 945	15 930
1967-1968	1993-2007	26	17 688	15 357
1969-1970	1995-2007	26	16 057	13 476
1971-1972	1997-2007	26	15 095	12 209
1973-1974	1999-2007	26	14 248	11 343
1975-1976	2001-2007	26	13 482	9 548
Total			320 729	314 916

Table 1: Cohorts included in the analysis. Note: Age is defined by the older of the two birth cohorts.

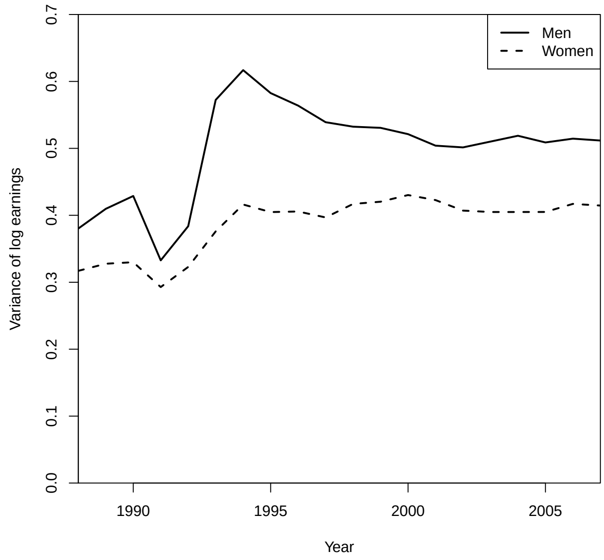
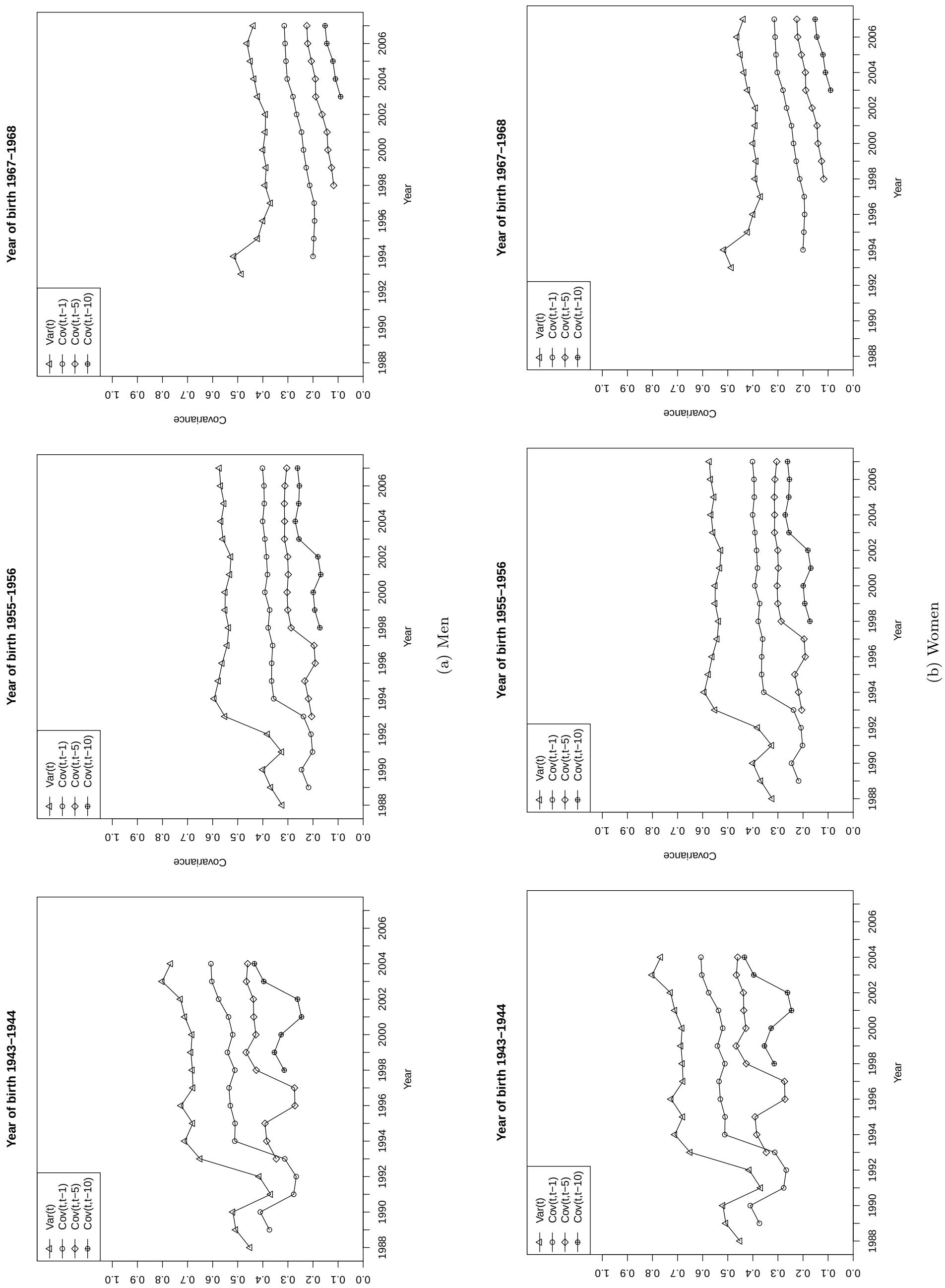


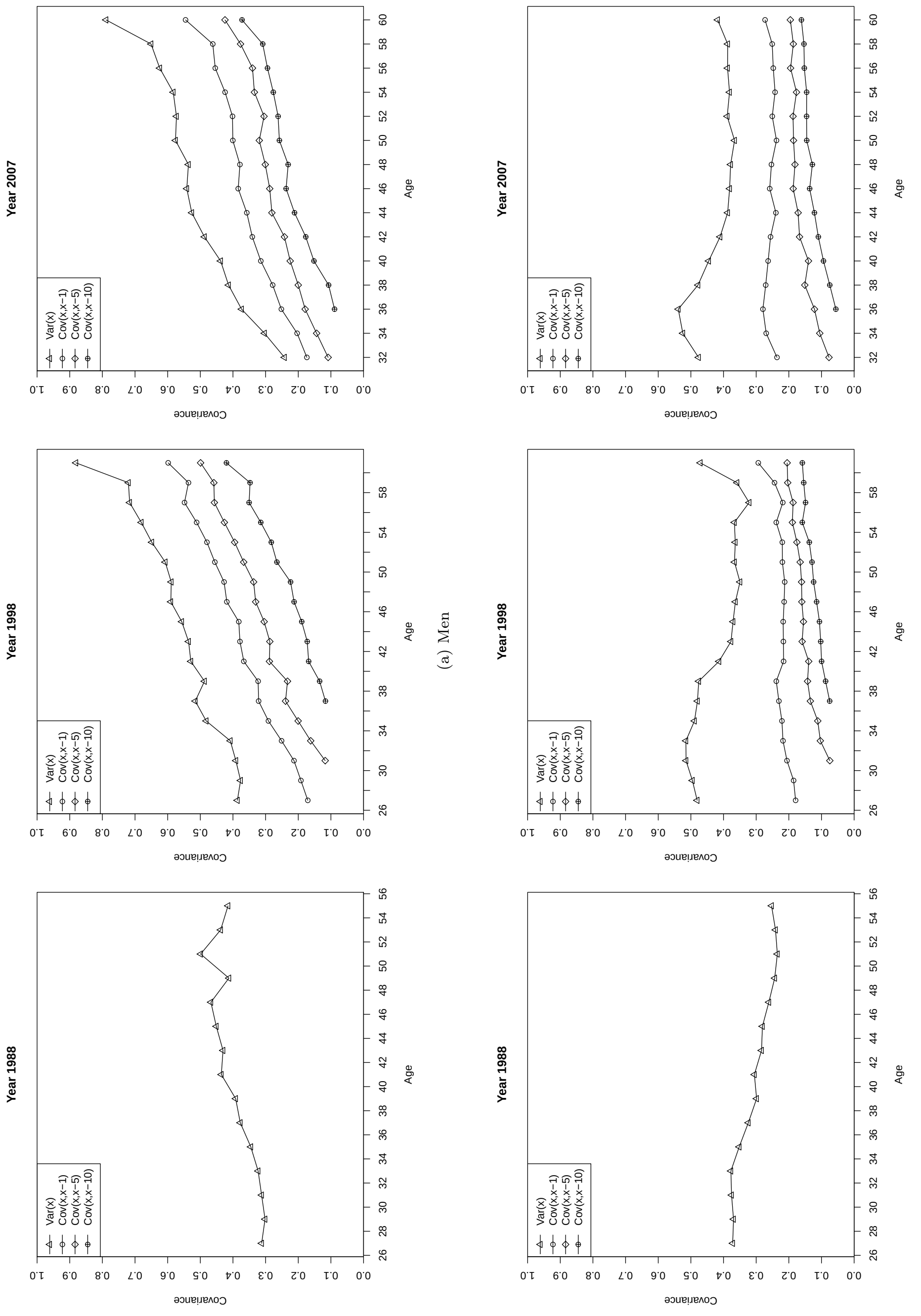
Fig. 1: Yearly earnings inequality (measured by variance of log earnings of workers) of men (solid line) and women (dashed line)



(a) Men

(b) Women

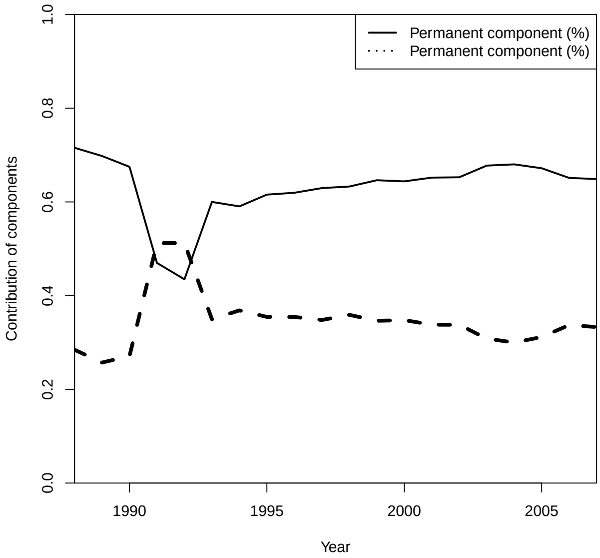
Fig. 2: Autocovariances of yearly log earnings for selected cohorts



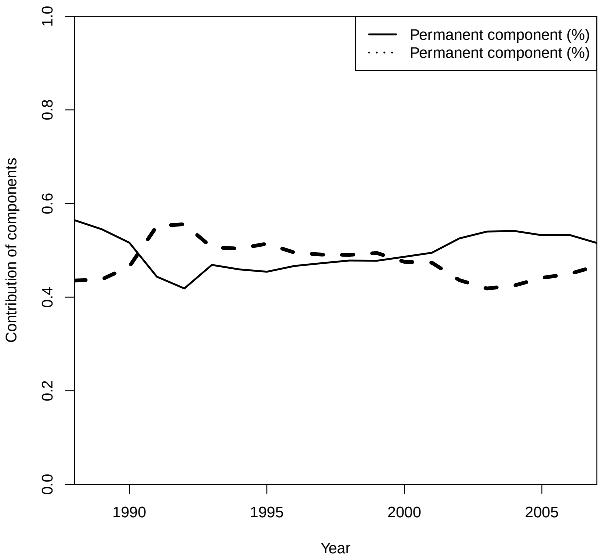
(a) Men

(b) Women

Fig. 3: Autocovariances of log yearly earnings for selected years



(a) Men



(b) Women

Fig. 4: Decomposition of the variance of log earnings among men and women measured in percentages. Predicted variance is calculated as the sum of the persistent and transitory components

	Men Parameter	S.E.	**Women** Parameter	S.E.
p_{1988}	1		1	
p_{1989}	1.017	0.010	0.983	0.011
p_{1990}	1.014	0.009	0.944	0.010
p_{1991}	0.737	0.019	0.811	0.009
p_{1992}	0.754	0.019	0.813	0.008
p_{1993}	1.071	0.016	0.913	0.012
p_{1994}	1.102	0.016	0.937	0.009
p_{1995}	1.094	0.018	0.907	0.012
p_{1996}	1.079	0.016	0.908	0.016
p_{1997}	1.063	0.014	0.892	0.015
p_{1998}	1.058	0.017	0.907	0.016
p_{1999}	1.067	0.020	0.898	0.016
p_{2000}	1.054	0.016	0.904	0.013
p_{2001}	1.043	0.014	0.894	0.015
p_{2002}	1.043	0.015	0.893	0.015
p_{2003}	1.063	0.014	0.889	0.012
p_{2004}	1.067	0.013	0.876	0.011
p_{2005}	1.044	0.013	0.855	0.012
p_{2006}	1.028	0.013	0.854	0.013
p_{2007}	1.017	0.015	0.825	0.012
$q_{1933-1934}$	0.977	0.016	0.777	0.010
$q_{1935-1936}$	1.054	0.012	0.846	0.007
$q_{1937-1938}$	1.076	0.009	0.847	0.005
$q_{1939-1940}$	1.005	0.007	0.876	0.004
$q_{1941-1942}$	1.049	0.005	0.904	0.003
$q_{1943-1944}$	1.047	0.004	0.940	0.003
$q_{1945-1946}$	1.033	0.003	0.938	0.002
$q_{1947-1948}$	1.034	0.002	0.958	0.001
$q_{1949-1950}$	1.004	0.001	0.971	0.001
$q_{1951-1952}$	1		1	
$q_{1953-1954}$	1.000	0.001	1.014	0.001
$q_{1955-1956}$	1.004	0.002	1.059	0.002
$q_{1957-1958}$	1.026	0.004	1.097	0.003
$q_{1959-1960}$	1.010	0.005	1.148	0.005
$q_{1961-1962}$	1.037	0.007	1.187	0.007
$q_{1963-1964}$	1.018	0.009	1.192	0.010
$q_{1965-1966}$	0.992	0.011	1.217	0.012
$q_{1967-1968}$	0.952	0.013	1.223	0.015
$q_{1969-1970}$	0.922	0.015	1.267	0.018
$q_{1971-1972}$	0.907	0.017	1.260	0.023
$q_{1973-1974}$	0.863	0.020	1.294	0.031
$q_{1975-1976}$	0.774	0.026	1.129	0.050
σ_a^2	0.156	0.003	0.093	0.002
σ_b^2	$1.5 * 10^{-5}$	$8 * 10^{-6}$	$2.9 * 10^{-5}$	$6 * 10^{-6}$
σ_{ab}	0.004	$2 * 10^{-4}$	0.004	$1 * 10^{-4}$

Table 2: Estimated parameters of permanent component of earnings.

	Men Parameter	S.E.	Women Parameter	S.E.
λ_{1988}	(unrestricted)		(unrestricted)	
λ_{1989}	1		1	
λ_{1990}	1.057	0.018	1.043	0.019
λ_{1991}	1.279	0.067	1.086	0.024
λ_{1992}	1.380	0.066	1.156	0.024
λ_{1993}	1.384	0.050	1.207	0.028
λ_{1994}	1.487	0.060	1.273	0.031
λ_{1995}	1.413	0.057	1.276	0.032
λ_{1996}	1.395	0.063	1.254	0.036
λ_{1997}	1.349	0.065	1.244	0.037
λ_{1998}	1.366	0.079	1.277	0.039
λ_{1999}	1.335	0.068	1.297	0.043
λ_{2000}	1.331	0.075	1.290	0.047
λ_{2001}	1.290	0.072	1.287	0.047
λ_{2002}	1.288	0.073	1.214	0.051
λ_{2003}	1.243	0.071	1.203	0.062
λ_{2004}	1.239	0.076	1.230	0.057
λ_{2005}	1.253	0.071	1.271	0.054
λ_{2006}	1.312	0.080	1.319	0.050
λ_{2007}	1.302	0.073	1.358	0.045
$\sigma^2_{1933-1934}$	0.025	0.009	0.058	0.005
$\sigma^2_{1935-1936}$	0.003	0.009	0.023	0.004
$\sigma^2_{1937-1938}$	0.067	0.009	0.029	0.004
$\sigma^2_{1939-1940}$	0.054	0.007	0.035	0.004
$\sigma^2_{1941-1942}$	0.097	0.008	0.053	0.004
$\sigma^2_{1943-1944}$	0.101	0.007	0.071	0.004
$\sigma^2_{1945-1946}$	0.109	0.007	0.088	0.004
$\sigma^2_{1947-1948}$	0.132	0.006	0.117	0.004
$\sigma^2_{1949-1950}$	0.123	0.006	0.120	0.003
$\sigma^2_{1951-1952}$	0.128	0.005	0.150	0.003
$\sigma^2_{1953-1954}$	0.113	0.005	0.188	0.003
$\sigma^2_{1955-1956}$	0.106	0.004	0.216	0.003
$\sigma^2_{1957-1958}$	0.103	0.004	0.220	0.003
$\sigma^2_{1959-1960}$	0.117	0.004	0.218	0.003
$\sigma^2_{1961-1962}$	0.135	0.003	0.230	0.002
$\sigma^2_{1963-1964}$	0.141	0.004	0.227	0.004
$\sigma^2_{1965-1966}$	0.210	0.004	0.254	0.002
$\sigma^2_{1967-1968}$	0.323	0.005	0.357	0.002
$\sigma^2_{1969-1970}$	0.281	0.003	0.361	0.002
$\sigma^2_{1971-1972}$	0.246	0.003	0.318	0.003
$\sigma^2_{1973-1974}$	0.255	0.004	0.362	0.004
$\sigma^2_{1975-1976}$	0.205	0.005	0.318	0.007
ρ	0.223	0.012	0.226	0.009
γ_0	0.112	0.013	0.249	0.015
γ_1	-0.001	0.001	-0.011	0.001
γ_2	$9 * 10^{-6}$	$2.7 * 10^{-5}$	$2 * 10^{-4}$	$3 * 10^{-5}$

Table 3: Estimated parameters of transitory variance of earnings.

	Random walk + AR(1) Parameter	S.E.	Random growth + AR(1) Parameter	S.E.	Random growth + ARMA(1,1) Parameter	S.E.
Men						
Permanent component						
σ_α	0.209	0.008	0.156	0.003	0.157	0.003
σ_β	$3.7 * 10^{-5}$	$1.5 * 10^{-5}$	$1.5 * 10^{-5}$	$8 * 10^{-6}$	$1.4 * 10^{-5}$	$8 * 10^{-6}$
$\sigma_{\alpha\beta}$	(restricted to 0)		0.004	$2 * 10^{-4}$	0.004	0.000
Transitory component						
ρ	0.87	0.016	0.223	0.012	0.258	0.042
δ	(restricted to 0)		(restricted to 0)		-0.032	0.041
γ_0	0.084	0.007	0.112	0.013	0.112	0.013
γ_1	0.003	0.001	-0.001	0.001	-0.001	0.001
γ_2	$-9.3 * 10^{-5}$	$2.3 * 10^{-5}$	$9 * 10^{-6}$	$2.7 * 10^{-5}$	$9 * 10^{-6}$	$2.7 * 10^{-5}$
Mean Square Error	1.81		1.29		1.29	
Women						
Permanent component						
σ_α	0.110	0.006	0.093	0.002	0.092	0.002
σ_β	$7 * 10^{-6}$	$1.1 * 10^{-5}$	$8 * 10^{-6}$	$2.9 * 10^{-5}$	$3.1 * 10^{-5}$	$7 * 10^{-6}$
$\sigma_{\alpha\beta}$	(restricted to 0)		0.004	$1 * 10^{-4}$	0.004	$1 * 10^{-4}$
Transitory component						
ρ	0.634	0.08	0.226	0.009	0.154	0.019
δ	(restricted to 0)		(restricted to 0)		0.069	0.026
γ_0	0.228	0.011	0.249	0.015	0.251	0.014
γ_1	-0.007	0.001	-0.011	0.001	-0.011	0.001
γ_2	$5.5 * 10^{-5}$	$3.2 * 10^{-5}$	$2 * 10^{-4}$	$3 * 10^{-5}$	$2 * 10^{-4}$	$3 * 10^{-5}$
MSE	1.42		0.69		0.68	

Table 4: Comparison of model specifications.

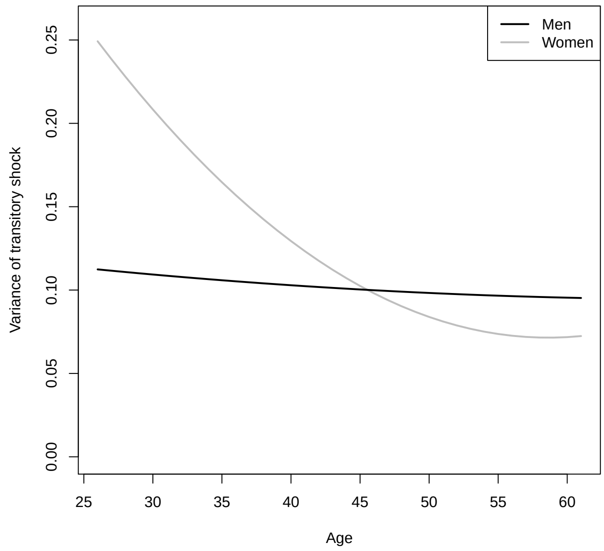
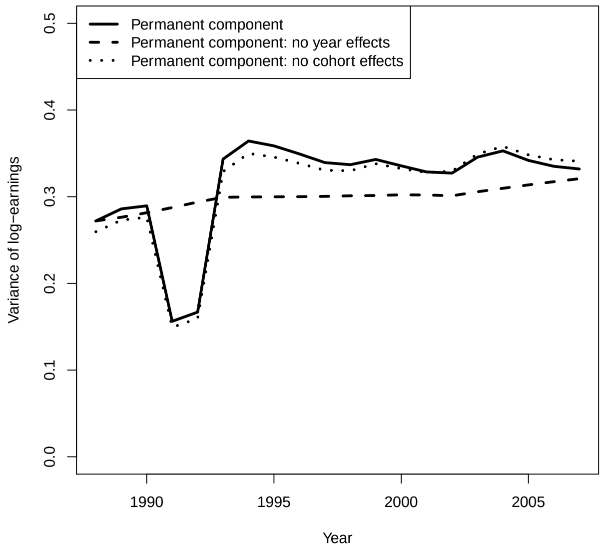
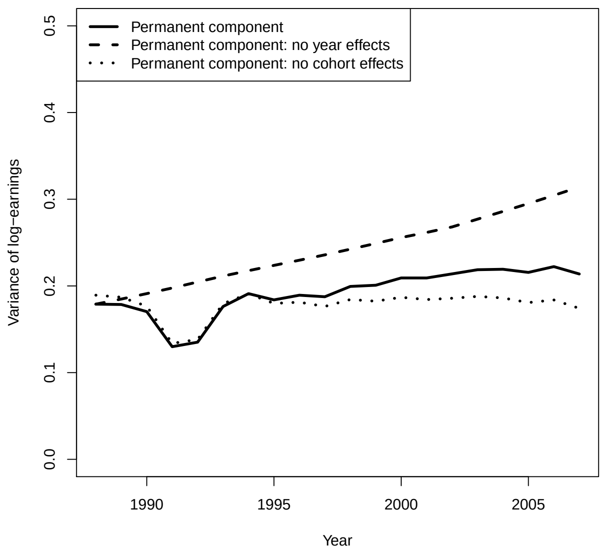


Fig. 5: Variance of the transitory shocks of men and women as a function of age

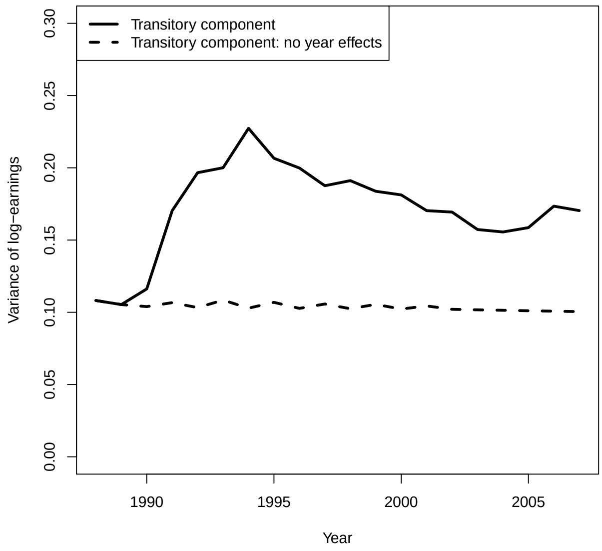


(a) Men

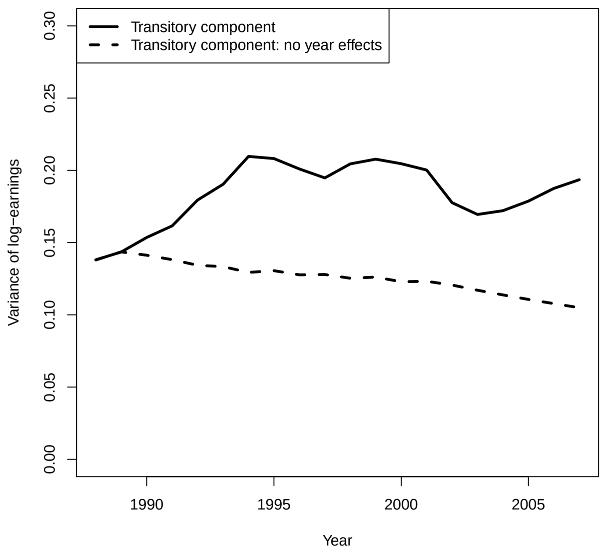


(b) Women

Fig. 6: The effects of eliminating year and cohort loadings on permanent earnings inequality for men and women



(a) Men



(b) Women

Fig. 7: The effects of eliminating year loadings on transitory earnings inequality for men and women