

Gustavsson, Magnus

**Working Paper**

## Trends in the Transitory Variance of Earnings: Evidence from Sweden 1960-1990 and a Comparison with the United States

Working Paper, No. 2004:11

**Provided in Cooperation with:**

Department of Economics, Uppsala University

*Suggested Citation:* Gustavsson, Magnus (2004) : Trends in the Transitory Variance of Earnings: Evidence from Sweden 1960-1990 and a Comparison with the United States, Working Paper, No. 2004:11, Uppsala University, Department of Economics, Uppsala

This Version is available at:

<https://hdl.handle.net/10419/82660>

**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.*



UPPSALA  
UNIVERSITET

Working Paper 2004:11

*Department of Economics*

Trends in the Transitory Variance of  
Earnings: Evidence from Sweden  
1960-1990 and a Comparison with  
the United States

Magnus Gustavsson

Department of Economics  
Uppsala University  
P.O. Box 513  
SE-751 20 Uppsala  
Sweden  
Fax: +46 18 471 14 78

Working paper 2004:11  
August 2004  
ISSN 0284-2904

TRENDS IN THE TRANSITORY VARIANCE OF EARNINGS:  
EVIDENCE FROM SWEDEN **1960-1990** AND A COMPARISON WITH THE UNITED STATES

MAGNUS GUSTAVSSON

Papers in the Working Paper Series are published  
on internet in PDF formats.  
Download from <http://www.nek.uu.se>  
or from S-WoPEC <http://swopec.hhs.se/uunewp/>

**Trends in the Transitory Variance of Earnings:  
Evidence from Sweden 1960-1990 and a Comparison with the United States\***

Magnus Gustavsson\*

June 8, 2004

**Abstract**

I decompose the cross-sectional variance of male annual earnings in Sweden between 1960 and 1990 into permanent and transitory components. The transitory variance increased until the early 1970s, declined during the remainder of the decade and then rose again during the second half of the 1980s. The permanent variance declined over the whole sample period but its decrease was much more rapid up until the early 1980s than afterwards. Comparing the results for the transitory variance with evidence from the U.S. reveals sharp differences. Most notably, the transitory variance of U.S. earnings rose sharply from the mid 1970s to the mid 1980s. An important explanation for these dissimilarities appears to be labor market institutions. In particular, it is likely that centralized solidarity bargaining in Sweden imposed constraints on earnings instability during the 1970s and early 1980s.

*JEL classification: C33, D31, J39*

*Keywords: Earnings dynamics, Earnings inequality, Panel data*

---

\* I am grateful to Meredith Beechey, Per-Anders Edin, Patrik Hesselius, Bertil Holmlund, Stephen Jenkins, Markus Jäntti, and seminar participants at Uppsala University as well as participants at the 2003 IZA Summer School for discussions and helpful comments.

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

## 1. Introduction

The narrowing of earnings differentials in most advanced countries during the 1960s and 1970s, and subsequent widening during the 1980s, is well documented. However, countries differed somewhat in the timing and magnitude of these changes, with two extreme observations being the U.S. and Sweden. Earnings differentials in the U.S. were fairly stable during the 1960s and increased slightly in the 1970s. In contrast, Sweden experienced precipitous pay compression from the late 1960s through the 1970s, which resulted in one of the most compressed wage structures visible among advanced countries. Earnings differentials increased in both countries during the 1980s but the increase in the U.S. was rapid and powerful whereas Sweden, by comparison, only experienced minor changes.<sup>1</sup>

Most explanations put forward to address the rapid increase in U.S. earnings inequality presume that it can be attributed to an increase in the dispersion of permanent earnings, such as an increase in the price of skill. The most popular explanations include skill-biased technological change and increased trade with developing countries. However, several studies (Gottschalk and Moffitt, 1994; Haider, 2001; Moffitt and Gottschalk, 2002) have shown that between a third and a half of the rise in the variance of U.S. earnings has in fact been due to increased earnings instability. This increase in the transitory variance of earnings should have rather different explanations than sole changes in the price of skill.

Little is known about the forces driving the increase in earnings instability in the U.S. labor market. Since this hinders the understanding of changes in earnings inequality, Katz and Autor (1999 p.1497) state that “An important agenda for future work is to attempt to examine the extent to which patterns of changes in transitory earnings variability are related to changes in technology, organizational and personnel practices, exposure to international competition, changes in domestic product market competition, and changes in unionization and other labor

---

<sup>1</sup> See Katz and Autor (1999) for wages and annual earnings in the U.S with an international comparison. For changes in the Swedish wage structure, see Edin and Holmlund (1995). For annual earnings inequality in Sweden since 1968, see e.g. Edin and Fredriksson (2000).

market institutions.”. This paper responds to this statement by offering a comparison of the evolution of the transitory variance of earnings in the Swedish and the U.S. labor markets. By comparing two countries with markedly different labor market institutions and markedly different earnings inequality experiences, more light can be shed on the forces underlying changes in earnings instability.

Sweden is also, by itself, well suited to a study of how changes in the functioning of the labor market co-vary with changes in the transitory variance of earnings. In particular, several sharp institutional changes have occurred in Sweden between 1960 and 1990 which seem to have affected wage dispersion (see Hibbs, 1990; Edin and Holmlund, 1995; Edin and Topel, 1997; Hibbs and Locking 2000). One notable turning point was the 1983 breakdown of centralized wage bargaining, present since the mid 1950s, and the subsequent shift to a system of industry-wide bargaining.

Another important and interesting period in the Swedish labor market was the powerful implementation of the “solidarity wage policy” pursued by the major blue-collar union, Landsorganisationen (LO), whose heyday coincides with the great wage compression beginning in the late 1960s. The policy was gradually brought into practice during the 1950s and had as its aim “equal pay for equal work”, meaning that wage differentials were to be independent of differences in profitability between regions, industries or firms. However, from the late 1960s to the breakdown of centralized bargaining in 1983, the policy also successfully invoked equalization of wages between all kinds of jobs by giving especially large pay increases to the lowest paid workers.

Relatively little is known about permanent or transitory earnings inequality in Sweden prior to 1968. Previous studies examining either earnings inequality or earnings dynamics before 1968 have either used data for a specific group within the labor market for which data has been available, or been forced to use small samples and a broader income measure which

has included capital income and realization of capital gains.<sup>2</sup> In fact, very little is known about changes in the transitory variance of earnings during the whole period 1960-90. This scarcity of information has limited the understanding of the Swedish labor market from a long-term perspective.

It is crucial to know whether changes in the cross-sectional variance stem from changes in the dispersion of permanent or transitory earnings as they have different implications for inequality. An increase in the dispersion of permanent earnings leads to greater earnings differentials in both the short- and long-run. An increase in the dispersion of transitory earnings, on the other hand, increases earnings differentials as measured in a single year but there is no change in earnings differentials when measured over several years.

In this paper I fill part of the knowledge gap by offering a long-run perspective on the Swedish labor market using previously unexploited longitudinal earnings data for 1960-67 together with longitudinal earnings data for 1968-90, resulting in an earnings panel spanning 31 years. Based on these data, I am able to estimate how the permanent and transitory components of annual earnings inequality have evolved over the three decades spanning 1960 to 1990.

Ideally, one should also study wage dynamics over the same period. However, Gustavsson (2004) finds that the evolution of the transitory and permanent components of wage inequality has been similar to that of annual earnings inequality in Sweden between 1992 and 1999, despite large swings in unemployment during this period. The results in Dickens (2000) for the U.K. 1975-95 and in Haider (2001) for the U.S. 1967-90 indicate a similar conclusion. In Sweden from 1968 onwards, changes in permanent wage inequality are also quite well documented. This is important since it, to some extent, makes it possible to

---

<sup>2</sup> These studies include Spånt (1979), Hibbs (1990), Björklund (1993) and Edin and Holmlund (1995).

map out the contribution from variation in hours worked and variation in wages to the variation in annual earnings.

Previewing the main results found in this paper, the transitory variance is found to have increased until the early 1970s, declined during the remainder of the decade and then risen again during the second half of the 1980s. The permanent variance declined over the whole sample period but its decrease was much more rapid up until the early 1980s than afterwards. A comparison of the results for the transitory variance in Sweden with that in the U.S. evidence reveals sharp differences. In the U.S., earnings instability increased noticeably from the mid 1970s to the mid 1980s, and the increase was much stronger than anything experienced in Sweden between 1960 and 1990. An important explanation for the differences between the United States and Sweden appears to be differences in labor market institutions. In particular, it is likely that centralized solidarity bargaining in Sweden imposed constraints on earnings instability during the 1970s and early 1980s.

The remainder of the paper is organized as follows. Section 2 describes the parametric model of permanent and transitory inequality used in the analysis. Section 3 documents the data sources and the data preparation. Section 4 contains results for the parametric model outlined in section 2. Section 5 discusses the findings for earnings instability in Sweden and compares them to what is known about developments in the U.S. labor market. The paper ends with concluding remarks.

## **2. Parametric models of earnings dynamics**

This section describes the parametric model that I use to decompose changes in cross-sectional inequality into permanent (long-run) and transitory (instability) components. To



provide intuition behind this model, I first describe a more basic model of earnings dynamics.<sup>3</sup>

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

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

expresses  $Y_{ibt}$  as the cohort specific mean  $\mu_{bt}$  in year  $t$  plus an individual specific deviation  $y_{ibt}$  from that mean. A stripped down model for  $y_{ibt}$  is

$$(2) \quad y_{ibt} = p_t u_{ib} + \lambda_t \varepsilon_{ibt}.$$

In equation (2), the variable  $u_{ib}$  and its year-specific factor loading,  $p_t$ , capture permanent, or persistent, relative earnings. The variable  $u_{ib}$  can be thought of as capturing all individual characteristics that matter for permanent relative earnings and  $p_t$  as reflecting the price of these characteristics. The variable  $\varepsilon_{ibt}$  and its factor loading,  $\lambda_t$ , capture stochastic transitory deviations from permanent earnings. The transitory component is serially uncorrelated with mean zero, and  $u_{ib}$  and  $\varepsilon_{ibt}$  are uncorrelated with variances  $\sigma_u^2$  and  $\sigma_\varepsilon^2$  respectively. With these assumptions, the variance of log (relative) earnings is

$$(3) \quad \text{Var}(y_{ibt}) = p_t^2 \sigma_u^2 + \lambda_t^2 \sigma_\varepsilon^2,$$

and the auto-covariance between year  $t$  and  $t-s$  is

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

Equation (3) shows that an increase in either factor loading generates increased cross-sectional earnings dispersion. The character of the change depends crucially, however, on which of the factor loadings that changes. A persistent rise in  $p_t$  increases long-run inequality as the relative labor market advantage of workers with chronically high earnings is enhanced. An increase in  $\lambda_t$  without any change in  $p_t$  generates increased cross-sectional earnings

---

<sup>3</sup> The description of this simpler model draws heavily on Baker and Solon (2003)

dispersion by raising year-to-year earnings instability but with no change in long-run inequality.

Changes in the factor loadings are closely related to changes in measures of earnings mobility, i.e. changes in the rate at which individuals shift positions in the earnings distribution (transition across quantiles of the earnings distribution). Increases in  $p_t$  without a change in  $\lambda_t$  will cause the auto-covariances to grow in greater proportion than the variances, so auto-correlations increase. In contrast, increases in  $\lambda_t$  without a change in  $p_t$  will only increase the variances, so auto-correlations decrease. Equal proportional increases in the two factor loadings will leave mobility rates unchanged, even though individual earnings instability will be increased. Changes in auto-correlations thus identify changes in the ratio of persistent to transitory earnings inequality.<sup>4</sup>

The model in equation (2), although intuitive, is likely to be too restrictive to adequately capture changes in permanent and transitory inequality. First, over their life cycle, most workers experience changes in the personal characteristics that determine permanent earnings. Events like education and training, changes of employer and unemployment spells are all likely to affect permanent earnings. Since individuals have different experiences of these events, persistent inequality among individuals in the same cohort is likely to change over the life cycle. Second, and less often recognized, there might be important cohort differences in permanent inequality.<sup>5</sup> In the sample employed in this paper, the oldest cohort is born in 1912 and the youngest in 1959. These two cohorts were born into very different societies and participated in sufficiently different education systems that the magnitude of permanent earnings inequality may be affected. Third, most previous studies have found the transitory component to be serially correlated. Fourth, and seldom recognized, the magnitude of

---

<sup>4</sup> A more detailed discussion of the connection to earnings mobility can be found in Moffitt and Gottschalk (1995).

<sup>5</sup> Three studies that allow for this are Cappellari (2002), Baker and Solon (2003) and Ramos (2003).

transitory shocks probably varies with age.<sup>6</sup> Younger individuals might have less stable careers (and lives). Last but not least, there may exist cohort differences in transitory inequality which reflect the different environments cohorts have lived through.

To allow for all these aspects of earnings dynamics, a much more complex model is called for. After much experimentation, the following generalization of equation (2) is found to best fit the data:<sup>7</sup>

$$(5) \quad y_{ibt} = p_t q_b u_{ia} + \varepsilon_{ibt},$$

where

$$(6) \quad u_{ia} = u_{i,a-1} + r_{ia},$$

$$(7) \quad \varepsilon_{ibt} = \rho_t \varepsilon_{ib,t-1} + \delta_t \lambda_{t-1} v_{ib,t-1} + \lambda_t v_{ibt},$$

and

$$(8) \quad \text{Var}(v_{ibt}) = \gamma_0 + \gamma_1 a + \gamma_2 a^2,$$

where  $a = t - b - 27$ , i.e. years since age 27 (by definition the lowest age in my sample).

Beginning with the permanent component,  $q_b$  in equation (5) is a cohort-specific parameter that allows the dispersion of permanent earnings to differ between cohorts.<sup>8</sup>

Equations (5) and (6) describe the persistent component as a random walk in age where

$r_{i,a} \sim \text{iid}(0, \sigma_{r,a}^2)$  is the innovation at each age.<sup>9</sup> In previous research it is common to force the

innovation variance  $\sigma_{r,a}^2$  to be the same over the whole life cycle but because of my rich data,

---

<sup>6</sup> Two studies that allows for this are Gustavsson (2004) and Baker and Solon (2003).

<sup>7</sup> Some models produce negative estimates of some of the variances or show other clear signs of being over-specified – these models are naturally discarded. All the other models produce the same main pattern of changes in the permanent and transitory variances of log earnings. The model presented here is chosen because it contains the most relevant information about individual earnings dynamics. I have not applied Newey's (1985) specification test as the previous literature on earnings dynamics show that this test always (at least in the studies that I am aware of) rejects the hypothesis that the right model is specified. The drawbacks with this test for assessing the goodness of fit of models of earnings dynamics are further discussed in Baker (1997) and Baker and Solon (2003). A general critic of tests such as that proposed in Newey (1985) is also found in Leamer (1983).

<sup>8</sup> The terms “permanent” and “transitory” are questionable. However, to be consistent with previous studies, I use the term permanent for the non-mean reverting component and transitory for the mean reverting component.

<sup>9</sup> A random walk specification is also used, among others, in Dickens (2000) and Moffitt and Gottschalk (2002).

I do not have to impose this restriction.<sup>10</sup> Rather, the innovation variance is only restricted to be the same for two ages at a time, i.e. individuals aged 28 and 29 have one variance, those aged 29 and 30 have another variance, and so forth.<sup>11</sup> I also estimate the variance of an initial permanent shock, as of age 27, denoted  $\sigma_u^2$ . This specification of the permanent component recognizes that younger individuals are more likely to be engaged in job shopping and have no job securing tenure and therefore are likely to have larger permanent shocks to earnings.

For the transitory component, equation (7) models deviations from permanent earnings as a first order autoregressive moving average process with year-specific factor loadings,  $\lambda_t$ , on the innovation  $v_{ibt}$ . This specification implies that luck follows luck in the sense that a positive transitory shock dies out gradually. Equation (7) also allows the ARMA-parameters,  $\rho_t$  and  $\delta_t$ , to vary over time. More specifically, they are allowed to differ between, but not within, the periods 1960-67 and 1968-90; the reasons for this are discussed in the data section.

Equation (7) shows that transitory earnings for an individual in cohort  $b$  in year  $t$  is a function of transitory earnings in year  $t-1$ , which in turn is a function of transitory earnings in year  $t-2$ , and so on. This auto-regressive process induces a recursive structure in the moments. If one traces the recursion back to a cohort's first sample year, this raises the question of what the value of the transitory variance is in that 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 and Solon (2003) and estimate separate initial transitory variances for each cohort. This recognizes that earnings instability varies across cohorts because they are at different stages of the life cycle and have

---

<sup>10</sup> One exception is Dickens (2000).

<sup>11</sup> I have also estimated models where the variance is permitted to differ across all ages but have found that this does not improve the fit of the model noticeably.

lived through different times.<sup>12</sup> Consistent with separate initial variances, equation (8) allows the variance of  $v_{ibt}$  to be a quadratic function of age.

Although the model in equations (5)-(8) is more complex than that in equation (2), the intuition from the simpler model still holds. An increase in the permanent variance preserves the order of individuals in the earnings distribution but spreads them out further, and this greater spread remains year after year. An increase in the transitory variance leads to more scrambling of workers' order in the annual earnings distribution, and the scrambling gets redone every year.

Even though the model in equations (5)-(8) still imposes a great deal of structure on individual earnings, it is much more flexible than the models used in the U.S. studies by Haider (2001) and Moffitt and Gottschalk (2002). These studies do not allow for cohort heterogeneity in the permanent and transitory components or age heteroskedasticity in the transitory variance. Their specification of the way in which permanent earnings varies with age is also more restrictive.<sup>13</sup>

The parameters of the model in equations (5)-(8) are estimated by applying the minimum distance estimator of Chamberlain (1984). Basically, the implied variances and auto-covariances of the model in equations (5)-(8) are fitted to the corresponding empirical moments in the data by non-linear least squares. Appendix A contains a description of the estimation procedure.

---

<sup>12</sup> One might argue that the transitory component should also be multiplied by a cohort specific parameter. However, when estimated, such a model shows clear signs of being over-specified as some of the variances of the transitory shocks become negative. I therefore restrict the cohort-effects to act through the initial transitory variance.

<sup>13</sup> Baker and Solon (2003) use Canadian annual earnings to test how their results are affected by the restrictions imposed in the U.S. studies. They find that the main conclusions regarding *changes* in the permanent and transitory component are unaffected, but that the relative size of the transitory component is somewhat overestimated. I have also experimented with these restrictions and generally reached similar conclusions. Of course, these results need not carry over to the U.S. context.

### **3. Data**

This section is divided into three sub-sections. The first describes the data source and the sample construction. The second contains an overview of annual earnings dispersion in Sweden during 1960-90. Part of the earnings data for 1960-1967 are top-coded; the last sub-section describes how I have adjusted these data.

#### 3.1 Data source and sample construction

The data come from the Swedish longitudinal database LINDA (Longitudinal Individual Data for Sweden), constructed to be cross-sectionally representative of the Swedish population each year from 1968 onwards (see Edin and Fredriksson, 2000). The dataset is large; it contains 3.35 percent of the Swedish population each year, amounting to around 300,000 individuals. All information is based on administrative registers, which confers several advantages compared to an analysis based on survey data. First, there is no outflow apart from death or migration, thus the data are free of the kind of sample attrition common in surveys. Second, the data are highly reliable; information from administrative registers is likely to be better than the recall of individuals. A drawback with the data is the limited information on an individual's labor market status. In particular, there is limited information on why an individual has very low or zero earnings.

Information on individual earnings is annually available for the LINDA samples from 1968 onwards in the Income Register, which is based on tax reports. For the 1960-67 period, I use the Pensionable Income Register. This register is available for the 1990 LINDA sample and contains information on annual earnings from 1960 onwards that form the basis for supplementary old-age pensions (“ATP-poäng”).<sup>14</sup>

---

<sup>14</sup> See Riksförsäkringsverket (1984) for detailed information on the calculation of earnings that form the basis for supplementary old-age pension.

As the purpose of the 1960-67 data is to inform on supplementary old-age pensions for the 1990 LINDA sample, it suffers from several shortcomings from the point of view of the current analysis. First, individuals working in the period 1960-67 must still be alive and living in Sweden in 1990 in order to be observed. This means that the earnings sample 1960-67 might not be random, as people with bad health probably are less likely to still be alive in 1990. Second, information on annual earnings is missing if an individual had any form of pension during a year, because these years did not qualify for supplementary old-age pension. Third, earnings have to be above a specific threshold to form the basis for supplementary old-age pension – the so-called “basic amount”. The basic amount was 4,200 SEK in 1960, which corresponds to 41,143 SEK in 2002. The basic amount is annually corrected for changes in consumer prices, why the real threshold is constant. Lastly, part of the earnings data are top-coded because earnings above 7.5 basic amounts did not form the basis for supplementary old age pension. Because the ceiling is constant in real terms and real wages increased during the 1960s, the percentage of earnings that are top-coded increases over time. Despite the drawbacks with the 1960-67 data, it constitutes a unique opportunity to gain insights into individual earnings dynamics in Sweden before 1968 – something that few previous studies have been able to touch on.

The measure of earnings used in the analysis contains earnings from all jobs, including self-employment, held by an individual during a year and is converted into 1990 real SEK using the consumer price index (“KPI”). During 1968-73 it is not possible to separate pensions from earnings. During 1960-73 it is also not possible to include sickness benefits in the measure of earnings. To make the measure consistent over time, sickness benefits are excluded from the earnings variable during all years. In the empirical analysis I find that changes in the sickness absence rate have no significant effect on the estimated permanent and transitory variances.

The analysis in this paper is based on men born in Sweden aged 26-53. This age range should ensure that individuals are old enough to have completed their education and young enough not to be considered for early retirement. The upper age of 53 hence limits the effects of missing information on annual earnings in the 1960-67 data for individuals receiving pension during part of a year and of not being able to separate pensions from earnings during 1968-73. It should also limit the potential sample selection in the 1960-67 data. The restriction to males in the study is because the large changes in female labor force participation during the sample period would confound an analysis of female earnings. The results would be confounded because I estimate a model of earnings, not of entry and exit; appending a model of entry and exit is beyond the scope of this paper. The same argument applies to the exclusion of foreign-born men, as there has been substantial immigration into Sweden during the sample period.

For the empirical analysis, I categorize individuals into two-year birth cohorts and follow them through time. I include cohorts who are between the age 26 and 53 for at least 6 years between 1960 and 1990. The youngest two-year cohort is aged 26 and 27 in 1985 (born in 1958/59), the next youngest is 26 and 27 in 1983 (born in 1956/57), and so on down to the oldest cohort, aged 47 and 48 in 1960 (born in 1912/13); in the estimations, the age of each two-year birth cohort is defined by the older of the two cohorts. Cohorts can be present in the sample between 6 and 27 years depending on their date of birth. This gives a total of 24 two-year birth cohorts. The second column of Table 1 contains for each year the resulting age intervals in the sample.

In order to maximize the sample used and not to transfer the potential sample selection effects present in 1960-67 to later data, I include all earnings observations larger than the basic amount for each individual over the period 1960-90, allowing individuals to re-enter the panel if they do exit. The restriction to earnings above the basic amount is to be able to get



**Table 1: Description of the sample**

| Year | Age   | # individuals in the sample | # individuals excluded due to earnings < basic amount | % top-coded in the sample |
|------|-------|-----------------------------|---|---------------------------|
| 1960 | 27-48 | 23,585                      | 2,069   | 4.64                      |
| 1961 | 26-49 | 25,962                      | 2,164   | 5.51                      |
| 1962 | 27-50 | 25,960                      | 2,166   | 6.44                      |
| 1963 | 26-51 | 28,503                      | 2,152   | 6.69                      |
| 1964 | 27-52 | 28,650                      | 2,005   | 8.74                      |
| 1965 | 26-53 | 31,433                      | 2,116   | 9.67                      |
| 1966 | 27-52 | 29,776                      | 2,039   | 11.32                     |
| 1967 | 26-53 | 32,409                      | 2,371   | 12.07                     |
| 1968 | 27-52 | 34,326                      | 4,272   | -                         |
| 1969 | 26-53 | 38,768                      | 3,681   | -                         |
| 1970 | 27-52 | 36,771                      | 2,467   | -                         |
| 1971 | 26-53 | 40,978                      | 2,686   | -                         |
| 1972 | 27-52 | 37,879                      | 2,625   | -                         |
| 1973 | 26-53 | 41,762                      | 2,842   | -                         |
| 1974 | 27-52 | 38,322                      | 2,689   | -                         |
| 1975 | 26-53 | 41,934                      | 2,954   | -                         |
| 1976 | 27-52 | 38,974                      | 2,611   | -                         |
| 1977 | 26-53 | 41,953                      | 3,096   | -                         |
| 1978 | 27-52 | 39,164                      | 2,794   | -                         |
| 1979 | 26-53 | 42,230                      | 3,214   | -                         |
| 1980 | 27-52 | 39,796                      | 2,812   | -                         |
| 1981 | 26-53 | 42,731                      | 3,326   | -                         |
| 1982 | 27-52 | 39,887                      | 3,324   | -                         |
| 1983 | 26-53 | 42,815                      | 3,727   | -                         |
| 1984 | 27-52 | 40,427                      | 3,423   | -                         |
| 1985 | 26-53 | 43,393                      | 3,829   | -                         |
| 1986 | 27-52 | 41,089                      | 3,547   | -                         |
| 1987 | 28-53 | 41,142                      | 3,344   | -                         |
| 1988 | 29-52 | 38,873                      | 3,036   | -                         |
| 1989 | 30-53 | 38,728                      | 3,075   | -                         |
| 1990 | 31-52 | 36,447                      | 3,014   | -                         |

comparable estimates of earnings inequality for the 1960-67 and 1968-90 periods; in the empirical analysis I investigate the effect on the estimates from this threshold.

It might be technically possible to use a sample in which all earnings below the basic amount are set to zero and then include all the zero earnings in the analysis. However, including these would combine the analysis of earnings with the dynamics associated with the extensive and intensive margins of working during a year, making the estimation results hard to interpret.

The end result of the sample construction is an unbalanced panel as some individuals die or migrate during the sample period and others do not have earnings above the basic amount during the whole sample period. The third and fourth columns of Table 1 contain for each year the total sample size and the number of individuals excluded because they have earnings below the basic amount. The last column also reports the percentage of earnings that are top-coded in the 1960-67 data. In total, the 1960-90 panel consists of 76,079 men and 1,144,667 individual-year observations. This should be compared to the sample sizes of 3,115 and 2,988 individuals used in the equivalent U.S. studies by Haider (2001) and Moffitt and Gottschalk (2002), respectively.

For the two-year birth cohorts, the sample sizes range between 1,572 individuals in 1960 for the cohort born in 1912/13, to 4,183 individuals in 1976 for the cohort born in 1944/45. Table 2 presents the sample size and the extent to which individuals disappear from the sample over time for selected two-year birth cohorts and years. For example, for the cohort born 1928/29, 2309 individuals are present in 1960. Of these, 98 percent are present in 1965, and 93 percent are still present in 1980. For the period 1960-67, individuals can only disappear from the sample due to earnings below the basic amount. For the years 1968-90, individuals who disappear after one to five years do so mainly because they have earnings below the basic amount, while complete absence from the LINDA database becomes relatively more important over time. Taking the cohort born in 1948/49 in Table 2 as an example, of the 3679 individuals present in this cohort in 1975, 106 are not present in 1976. Of these 106 individuals, 5 have disappeared from the LINDA database, 33 have zero earnings, and 68 have positive earnings less than the basic amount. In 1990, 271 of the 3679 individuals are not present. Of these 271 individuals, 96 have disappeared from the LINDA database, 83 have zero earnings, and 94 have positive earnings less than the basic amount.

**Table 2: Sample size and changes in sample composition for selected cohorts and years**

| Year                | Sample size | % of these present after 1 year | % of these present after 5 years | % of these present after 10 years | % of these present after 15 years | % of these present after 20 years | % of these present after 25 years |
|---------------------|-------------|---------------------------------|----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
| Cohort born 1918/19 |             |                                 |                                  |                                   |                                   |                                   |                                   |
| 1960                | 2095        | 99                              | 98                               | 96                                | -                                 | -                                 | -                                 |
| 1965                | 2117        | 98                              | 96                               | -                                 | -                                 | -                                 | -                                 |
| 1970                | 2956        | 97                              | -                                | -                                 | -                                 | -                                 | -                                 |
| Cohort born 1928/29 |             |                                 |                                  |                                   |                                   |                                   |                                   |
| 1960                | 2309        | 99                              | 98                               | 96                                | 96                                | 93                                | -                                 |
| 1965                | 2345        | 99                              | 96                               | 96                                | 93                                | -                                 | -                                 |
| 1970                | 2673        | 98                              | 95                               | 89                                | -                                 | -                                 | -                                 |
| 1975                | 2600        | 98                              | 93                               | -                                 | -                                 | -                                 | -                                 |
| 1980                | 2469        | 97                              | -                                | -                                 | -                                 | -                                 | -                                 |
| Cohort born 1938/39 |             |                                 |                                  |                                   |                                   |                                   |                                   |
| 1965                | 2676        | 97                              | 95                               | 95                                | 95                                | 93                                | 91                                |
| 1970                | 2825        | 97                              | 95                               | 95                                | 91                                | 87                                | -                                 |
| 1975                | 2824        | 98                              | 97                               | 93                                | 88                                | -                                 | -                                 |
| 1980                | 2811        | 98                              | 94                               | 89                                | -                                 | -                                 | -                                 |
| 1985                | 2706        | 98                              | 92                               | -                                 | -                                 | -                                 | -                                 |
| 1990                | 2592        | -                               | -                                | -                                 | -                                 | -                                 | -                                 |
| Cohort born 1948/49 |             |                                 |                                  |                                   |                                   |                                   |                                   |
| 1975                | 3679        | 97                              | 95                               | 93                                | 93                                | -                                 | -                                 |
| 1980                | 3720        | 97                              | 95                               | 94                                | -                                 | -                                 | -                                 |
| 1985                | 3639        | 98                              | 96                               | -                                 | -                                 | -                                 | -                                 |
| 1990                | 3646        | -                               | -                                | -                                 | -                                 | -                                 | -                                 |
| Cohort born 1958/59 |             |                                 |                                  |                                   |                                   |                                   |                                   |
| 1985                | 3061        | 96                              | 95                               | -                                 | -                                 | -                                 | -                                 |
| 1990                | 3194        | -                               | -                                | -                                 | -                                 | -                                 | -                                 |

Overall, the extent to which individuals disappear from my sample must be considered to be very low. For example, in the corresponding U.K. study by Dickens (2000) it is reported that around 50 percent in a given cohort are still present after 20 years.<sup>15</sup> This large difference is because he uses survey data where attrition, i.e. non-response, is an issue, whereas my data is based on registers. It is also worth pointing out that the corresponding studies for the U.S. by Haider (2002) and Moffitt and Gottschalk (2003) use data from the survey PSID where attrition is extensive (see Fitzgerald *et al.*, 1998).

The sample size of a cohort may increase over time as new individuals enter the panel. This occurs mainly between 1967 and 1968, as some individuals present in the 1968 sample

<sup>15</sup> Other equivalent studies do not report these numbers.

are not alive and living in Sweden in 1990 and therefore do not have recorded earnings between 1960 and 1967. The oldest two-year cohort present in both 1967 and 1968, born 1916/17, also has the largest increase in sample size; the sample consists of 1,920 individuals in 1967, but of 2,809 individuals in 1968.

To be able to estimate the parameters of the model in equations (5)-(8), I take each cohort separately and, after adjusting the top-coded earnings in the period 1960-67 (described below), estimate the variances and auto-covariances of log annual earnings. Computing auto-covariance matrices for each cohort results in a total of 4,188 distinct variance and auto-covariance elements.

Here, one drawback of the unbalanced sample is that the variance in year  $t$  is based on all individuals with positive earnings in that year, while the auto-covariance between year  $t$  and year  $t-s$  is based only on those individuals with positive earnings in both these years. One way to avoid this is to use a revolving balanced panel design. In this design, developed by Haider (2002), only individuals with consecutive years of positive earnings during their cohort's entire sample period are included. Despite the appealing features of this panel design, it is not an option here because the potential selection effects present for the 1960-67 data would extend into the rest of the sample years, as only those cohort members with positive earnings between 1960 and 1967 could be present in a cohort's later years. To complicate matters more, those cohorts that are not present between 1960 and 1967 would not be affected by this sample selection. A revolving balanced panel could also by itself induce sample selection effects because of the requirement of consecutive years of positive earnings, especially here where the panel spans 31 years and earnings above the basic amount would be required in each year. In the empirical analysis I investigate how the results are affected by the choice of panel design.

### 3.2 Overview of annual earnings dispersion in the sample

Before drawing conclusions about the 1960-67 period, it is important to see whether earnings inequality based on data from the Pensionable Income Register (henceforth PIR) and the Income Register (henceforth IR) follow the same pattern over time.<sup>16</sup> Figure 1 contains the P80-P20 ratio of earnings for my 1968-90 sample (based on the IR), where P80 is the 80<sup>th</sup> percentile of the earnings distribution. The figure also contains the P80-P20 ratio for 1960-88 based on earnings from the PIR with the same age interval as in my sample, which means that this series contains the P80-P20 ratio for my 1960-67 sample. The years 1989-90 of the PIR are excluded because more than 20 percent of the earnings are top-coded.

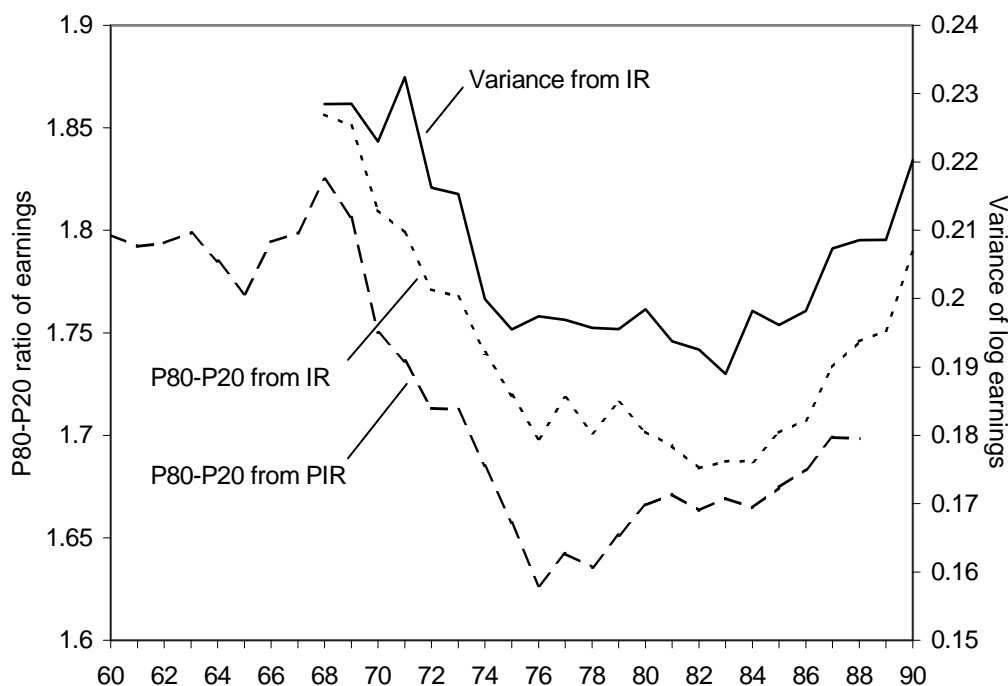
Figure 1 shows that the P80-P20 ratios based on earnings from the PIR and the IR follow the same general pattern over time. It is thus likely that they would display a similar pattern between 1960 and 1967 if data from IR were available for this period.<sup>17</sup> The magnitude of inequality is, however, lower for earnings based on the PIR because fewer individuals with low earnings are present in this register. This may be explained by the requirement of being alive in 1990 and the requirement of not having any pension during a year. From 1980 onwards, earnings that form the basis for supplementary old age pensions were also calculated for individuals with pension during a year. The effect of this is clearly seen in Figure 1 as the gap between the P80-P20 ratios decreases in 1980.

Figure 1 also contains the variance of log annual earnings for my 1968-90 sample (based on the IR). The picture produced by the variance is very similar to that from the P80-P20 ratios. There are therefore strong reasons to believe that if earnings from the IR were available for 1960-67, the variance of log earnings for 1960-67 would display the same evolution as the P80-P20 ratio these years.

---

<sup>16</sup> To be able to compare information from the registers I must make some adjustments to the measure of earnings from the PIR from 1974 onwards because several types of benefits then also qualify for supplementary old-age pension, such as unemployment benefits. To remove these benefits, information from the IR is used.

<sup>17</sup> The same is also true for the P50-P20 and P80-P50 ratios (available on request).



**Figure 1: Annual earnings inequality in Sweden based on data from the IR and the PIR.**

Note: Sample ages are used each year; see Table 1.

The second column of Table 1 shows that there is some variation in the age range in the sample, and this could affect the results in Figure 1. But a sample where the age interval each year is fixed at 26-53 produces almost identical results as in Figure 1. The main pattern is also the same for a sample aged 26-36 each year; the results in Figure 1 for 1960-67 are therefore not an artifact of sample selection. Hence, annual earnings dispersion was fairly constant between 1960 and 1968, decreased from the late 1960s to the mid 1970s, and increased after 1983.

Another study that investigates annual earnings inequality among men in Sweden between 1960 and 1967 is Spånt (1979). He also finds that inequality was fairly constant during these years. However, his results are not fully comparable to those reported here, as his measure of earnings also contains capital income and realization of capital gains.

Edin and Fredriksson (2000) use data from the IR in LINDA and report the standard deviation of log annual earnings for men between 1968 and 1990. They use the same earnings threshold as in this paper - the basic amount - but the definition of earnings differs slightly as

they include sickness benefits after 1973. The age interval also differs as they each year include men 20-64 years old. Nevertheless, their results are similar to those in Figure 1.

The pattern for 1968-90 produced by Figure 1 mimics the evolution of wage dispersion among men. Consistent with Figure 1, Edin and Holmlund (1995) report that wage dispersion rapidly decreased from the early to mid 1970s, was fairly constant until 1983, then slightly increased during the remainder of the period. Less is known about male wage dispersion between 1960 and 1967. Edin and Holmlund's (1995) paper contains the P75-P25 ratio of wages for white-collar male workers employed in the private sector between 1956 and 1990. The reported evolution of wage dispersion for this group mimics the pattern produced by Figure 1. To conclude, the evolution of earnings dispersion in my sample is very similar to what is known about Swedish wage dispersion.

### 3.3 Remedial measures for the 1960-67 top-coded earnings

The percentage of individuals with top-coded earnings steadily increases in the 1960-67 data (see the last column of Table 1). This will cause estimates of the variance to decrease over time. It will also affect changes in the auto-correlations of earnings, which identify changes in the ratio of permanent to transitory inequality. Unfortunately, there is no ideal way to adjust the top-coded earnings. For example, imputing the values by using some type of ordered probit would only capture the permanent part of earnings. Besides this, there is not enough information on personal characteristics to be able to perform this. Instead, I assume a distribution for earnings and, based on this distribution, approximate the top-coded observations with their estimated mean value. The estimation is done separately for each cohort and each year, that is, one set of parameters for the distribution is estimated for the

cohort born 1912/13 in 1960, and another set is estimated for this cohort in 1961, and so forth.<sup>18</sup>

Two different distributional assumptions are employed. As recommended by Cowell (2000), I assume that the upper tail of the earnings distribution for each cohort and year can be approximated by the Pareto distribution. As an alternative, I assume that earnings can be approximated by the lognormal distribution.<sup>19</sup>

The Pareto distribution can be written in linear form as

$$(9) \quad \ln N = k + a \ln Z,$$

where  $Z$  equals an earnings level and  $N$  is the number of individuals with an income above  $Z$ . Following Fichtenbaum and Shahidi (1988), the parameters  $k$  and  $a$  are consistently estimated by least squares for each cohort and year. Let  $q$  denote the lowest value of the open-ended interval of the Pareto distribution. The mean of this interval is estimated as

$$(10) \quad m(q, \infty) = [a/(a+1)]q.$$

As the Pareto distribution is generally considered only to be applicable for higher earnings, it is necessary to choose a cut-off point for the estimation of equation (9). The rule-of-thumb offered by the literature varies. For example, Kakwani (1980) suggests that 40 percent of all workers belong to the Pareto form, whereas Lydall (1968) proposes 15-20 percent. In the 1960-67 data, the highest share of top-coded earnings for a cohort is 16 percent. Based on the rules-of-thumb and to assure that enough non-censored earnings are included in the estimation of equation (9), I set the cut-off point to 25 percent each year. The uniform cut-off ensures an equal procedure over the years and allows comparisons over time.

To implement the lognormal distribution, the mean of log-earnings implied by the normal distribution for each cohort and year is estimated by standard maximum likelihood

---

<sup>18</sup> I make no attempts to impute values for earnings below the basic amount. As these earnings are truncated, I do not know whether an individual has zero earnings or not, and hence, whether an individual actually is in the labor market in a given year.

<sup>19</sup> The lognormal distribution is consistent with the assumption that the log of permanent earnings follows a random walk; see e.g. Neal and Rosen (2000).



methods. The top-coded log-earnings are replaced with that value at which the actual mean of all log-earnings for a given cohort and year is equal to that implied by the normal distribution.

Which distribution works best for the current purpose depends on how the estimates of the covariance structure of earnings are affected. In Appendix B I use artificially top-coded earnings to investigate how the variances and auto-correlations of log earnings are typically affected by the “mean-approximation method” based on the two distributions. For the estimated variances, the use of the lognormal distribution results in too low values but unlike for the Pareto distribution, the sign of the estimated changes over time are correct. For the estimated auto-correlation matrices, the use of the Pareto distribution works better than the lognormal distribution but the difference is small; both distributions generally produce auto-correlations slightly lower than the true values but with correct changes over time.

Overall, the analysis in Appendix B shows that the mean-approximations based on the lognormal distribution produce more accurate results, so this distribution is employed in the main analysis. However, the more correct auto-correlations produced by the Pareto distribution can further inform us about the *relative* importance of the permanent and the transitory components of inequality during 1960-67. As a check of the results based on the lognormal distribution, I therefore also present secondary estimations based on earnings where the Pareto distribution has been applied.

Although the lognormal distribution quite correctly captures changes in inequality, the estimated variances are smaller than the true values. But even if this were not the case, the P80-P20 ratios in Figure 1 indicate that the variances in the period 1960-67 would be smaller than for later periods because fewer individuals from the lower part of the earnings distribution are included during 1960-67. In order to be able to better compare the results based on the 1960-67 data to later periods, it is necessary to re-scale the variances produced by the 1960-67 earnings.

Let  $Var(Y_{ibt}^*)$  denote the correct variance of log earnings for cohort  $b$  in year  $t$ . Let  $Y_{ibt,norm}$  be the log of earnings from the PIR for individuals in cohort  $b$  where top-coded values have been approximated with their mean based on the lognormal distribution. Let  $k_b$  be a constant that is used to re-scale all the values of  $Y_{ibt,norm}$  for cohort  $b$  so that the variance of  $Y_{ibt}^*$  and  $Y_{ibt,norm}$  are equal. This constant should then fulfill the equality

$$(11) \quad \frac{\sum(k_b Y_{ibt,norm} - (\sum k_b Y_{ibt,norm})/n)^2}{(n-1)} = Var(Y_{ibt}^*),$$

which can be written

$$(12) \quad k_b = \sqrt{\frac{Var(Y_{ibt}^*)}{Var(Y_{ibt,norm})}}.$$

An estimate of  $k_b$  in 1968 is obtained by estimating  $Var(Y_{ibt}^*)$  with the non-censored, representative earnings data from the IR for 1968 and by estimating  $Var(Y_{ibt,norm})$  with data from the PIR for 1968.<sup>20</sup> All earnings between 1960 and 1967 for cohort  $b$  are then multiplied by their estimate of  $k_b$ . This cohort-specific procedure adjusts for the underestimation produced by the lognormal distribution but also recognizes that sample selection differs between cohorts. For example, the requirement of being alive in 1990 imposes more sample selection for the cohort born 1912/13 than for the cohort born 1940/41; consequently, older cohorts have higher estimates of  $k_b$ . The cohort specific auto-correlations, which identify the relative contribution from the persistent and transitory components, are unaffected by this re-scaling.

The re-scaling procedure assumes that the values of  $k_b$  are the same each year during 1960-68. In Appendix B I find that the mean-approximation based on the lognormal

---

<sup>20</sup> This means that I use information for the cohorts born 1912/13 and 1914/15 when they are 56 and 54 years old, respectively. For the sample ages, the total share of earnings top-coded in the 1968 PIR is 12.97 percent.

distribution sometimes produces larger underestimations of the variance when the percentage top-coded is higher. Because the percentage top-coded is higher in 1968 than in the early 1960s, the values of  $k_b$  are probably not the same for each year between 1960 and 1968, at least for some cohorts. Rather, the true value is probably lower in the early 1960s. However, the empirical analysis contains several sensitivity analyses and none of these indicate that this sort of bias affects the main results.

#### **4. Estimation results**

This section is divided into two sub-sections. The first presents the main estimation results for earnings dynamics between 1960 and 1990. I discuss the estimated parameters of the model in equations (5)-(8) and then use these to decompose the cross-sectional variance into permanent and transitory components. The second sub-section shows that the results in the first sub-section are robust to various changes in the sample construction.

##### 4.1 Earnings dynamics in Sweden

Table 3 contains the estimated parameters and variances of the model in equations (5)-(8). The first column of results refers to the permanent component and the second to the transitory component. Starting with the persistent component, first are the age specific variances of the innovation in the random walk. These are all significantly larger than zero. As the variance of a variable that follows a random walk is the sum of the variances of the innovation, the results imply that permanent inequality increases over the whole studied age-range. The innovation variances are largest at younger ages; the same result is also found for Sweden between 1990 and 1999 in Gustavsson (2004) and for the U.K. between 1975 and 1995 in Dickens (2000). The presence of larger permanent shocks at younger ages is consistent with matching models

**Table 3: Minimum distance estimates of earnings dynamics**

|  | Permanent component | Transitory component |
|--|---------------------|----------------------|
| <i>Random walk variances</i>                           |                     |                      |
| $\sigma_u^2$   | 0.0854 (0.0036)     |                      |
| $\sigma_{r,28-29}^2$                                   | 0.0167 (0.0010)     |                      |
| $\sigma_{r,30-31}^2$                                   | 0.0159 (0.0010)     |                      |
| $\sigma_{r,32-33}^2$                                   | 0.0157 (0.0010)     |                      |
| $\sigma_{r,34-35}^2$                                   | 0.0139 (0.0010)     |                      |
| $\sigma_{r,36-37}^2$                                   | 0.0131 (0.0011)     |                      |
| $\sigma_{r,38-39}^2$                                   | 0.0126 (0.0011)     |                      |
| $\sigma_{r,40-41}^2$                                   | 0.0138 (0.0012)     |                      |
| $\sigma_{r,42-43}^2$                                   | 0.0130 (0.0012)     |                      |
| $\sigma_{r,44-45}^2$                                   | 0.0121 (0.0013)     |                      |
| $\sigma_{r,46-47}^2$                                   | 0.0137 (0.0014)     |                      |
| $\sigma_{r,48-49}^2$                                   | 0.0128 (0.0016)     |                      |
| $\sigma_{r,50-51}^2$                                   | 0.0115 (0.0018)     |                      |
| $\sigma_{r,52-53}^2$                                   | 0.0105 (0.0028)     |                      |
| <i>ARMA-parameters</i>                                 |                     |                      |
| $\rho_{1960-67}$                                       |                     | 0.8190 (0.0209)      |
| $\delta_{1960-67}$                                     |                     | -0.5045 (0.0235)     |
| $\rho_{1968-90}$                                       |                     | 0.5726 (0.0165)      |
| $\delta_{1968-90}$                                     |                     | -0.2579 (0.0139)     |
| <i>Parameters of the variance of transitory shocks</i> |                     |                      |
| $\gamma_0$   |                     | 0.0526 (0.0027)      |
| $\gamma_1$   |                     | -7.9E-4 (2.4E-4)     |
| $\gamma_2$   |                     | 3.4E-5 (9.1E-6)      |
| <i>Birth year</i>                                      | $q_b$               | $\sigma_b^2$         |
| 1912/13  | 0.7799 (0.0247)     | 0.0548 (0.0095)      |
| 1914/15  | 0.7941 (0.0240)     | 0.0373 (0.0088)      |
| 1916/17  | 0.8361 (0.0242)     | 0.0572 (0.0088)      |
| 1918/19  | 0.8827 (0.0247)     | 0.0591 (0.0104)      |
| 1920/21  | 0.8368 (0.0225)     | 0.0544 (0.0089)      |
| 1922/23  | 0.8657 (0.0239)     | 0.0362 (0.0098)      |
| 1924/25  | 0.9134 (0.0237)     | 0.0670 (0.0103)      |
| 1926/27  | 0.8873 (0.0243)     | 0.0299 (0.0095)      |
| 1928/29  | 0.9153 (0.0246)     | 0.0620 (0.0101)      |
| 1930/31  | 0.9667 (0.0256)     | 0.0536 (0.0096)      |
| 1932/33  | 1.0038 (0.0265)     | 0.0442 (0.0089)      |
| 1934/35  | 1.0000              | 0.0507 (0.0077)      |
| 1936/37  | 0.9931 (0.0256)     | 0.0683 (0.0073)      |
| 1938/39  | 1.0505 (0.0271)     | 0.0835 (0.0068)      |
| 1940/41  | 1.0494 (0.0278)     | 0.0854 (0.0067)      |
| 1942/43  | 1.0795 (0.0271)     | 0.0789 (0.0063)      |
| 1944/45  | 1.1076 (0.0277)     | 0.0973 (0.0061)      |
| 1946/47  | 1.1309 (0.0304)     | 0.0903 (0.0059)      |
| 1948/49  | 1.1146 (0.0304)     | 0.0956 (0.0059)      |
| 1950/51  | 1.1521 (0.0334)     | 0.0852 (0.0060)      |
| 1952/53  | 1.1869 (0.0364)     | 0.1047 (0.0065)      |
| 1954/55  | 1.2306 (0.0381)     | 0.0975 (0.0061)      |
| 1956/57  | 1.3249 (0.0429)     | 0.0965 (0.0058)      |
| 1958/59  | 1.3657 (0.0466)     | 0.0969 (0.0059)      |

**Table 3: (cont.)**

| <i>Factor loadings</i> | Permanent component |          | Transitory component |          |
|------------------------|---------------------|----------|----------------------|----------|
|                        | $p_i$               |          | $\lambda_i$          |          |
| 1960                   | 1.0000              |          | 1.0000               |          |
| 1961                   | 0.9704              | (0.0063) | 1.0000               |          |
| 1962                   | 0.9533              | (0.0072) | 0.9847               | (0.0219) |
| 1963                   | 0.9210              | (0.0077) | 1.0165               | (0.0246) |
| 1964                   | 0.8937              | (0.0081) | 1.0251               | (0.0253) |
| 1965                   | 0.8653              | (0.0083) | 1.0047               | (0.0258) |
| 1966                   | 0.8640              | (0.0088) | 1.0365               | (0.0263) |
| 1967                   | 0.8590              | (0.0093) | 1.0540               | (0.0270) |
| 1968                   | 0.8455              | (0.0093) | 1.0552               | (0.0248) |
| 1969                   | 0.8426              | (0.0099) | 1.0432               | (0.0259) |
| 1970                   | 0.8170              | (0.0103) | 1.0660               | (0.0261) |
| 1971                   | 0.8127              | (0.0109) | 1.1355               | (0.0278) |
| 1972                   | 0.7898              | (0.0111) | 1.0421               | (0.0252) |
| 1973                   | 0.7732              | (0.0113) | 1.0452               | (0.0248) |
| 1974                   | 0.7343              | (0.0114) | 1.0328               | (0.0243) |
| 1975                   | 0.7129              | (0.0115) | 1.0176               | (0.0240) |
| 1976                   | 0.7198              | (0.0122) | 1.0201               | (0.0238) |
| 1977                   | 0.7100              | (0.0125) | 1.0060               | (0.0241) |
| 1978                   | 0.6995              | (0.0130) | 1.0103               | (0.0247) |
| 1979                   | 0.6911              | (0.0133) | 0.9819               | (0.0243) |
| 1980                   | 0.6849              | (0.0137) | 1.0245               | (0.0255) |
| 1981                   | 0.6616              | (0.0138) | 1.0182               | (0.0256) |
| 1982                   | 0.6509              | (0.0141) | 1.0142               | (0.0261) |
| 1983                   | 0.6340              | (0.0142) | 0.9872               | (0.0261) |
| 1984                   | 0.6387              | (0.0147) | 1.0330               | (0.0271) |
| 1985                   | 0.6332              | (0.0148) | 0.9884               | (0.0266) |
| 1986                   | 0.6297              | (0.0150) | 1.0104               | (0.0271) |
| 1987                   | 0.6312              | (0.0155) | 1.0335               | (0.0285) |
| 1988                   | 0.6116              | (0.0154) | 1.0950               | (0.0285) |
| 1989                   | 0.6012              | (0.0155) | 1.0823               | (0.0282) |
| 1990                   | 0.5996              | (0.0158) | 1.1564               | (0.0290) |

Note: The estimated model is outlined in equations (5)-(8). 4188 variances and covariances are used in the estimation, which in turn are based on 76,079 men and 1,144,667 individual-year observations. Heteroskedasticity and auto-correlation robust standard errors are in parentheses.

in which information about a worker's ability is revealed over time (see for example

Jovanovic, 1979).

An alternative specification for the permanent component is the random growth model

$\alpha_{ib} + \beta_{ib} a$ . The variance for this model is  $\sigma_\alpha^2 + 2a\sigma_{\alpha\beta} + \sigma_\beta^2 a^2$ , where  $\sigma_{\alpha\beta}$  is the covariance

between  $\alpha_{ib}$  and  $\beta_{ib}$ , which may be negative. This model implies that the variance of

permanent earnings follows a convex quadratic pattern over the life cycle. However, at a

given year, a graph of the relation between age and the longer lag auto-covariances, which

mainly reflect permanent inequality, indicates that permanent inequality follows a concave increasing pattern over the life cycle. A concave increasing pattern is also implied by the estimates in Table 3. The concave pattern is further confirmed by the significant *negative* estimate of  $\sigma_{\beta}^2$  and the positive estimate of  $\sigma_{\alpha\beta}$  that is obtained with the random growth model - this is independent of whether the 1960-67 data is included or not. Baker and Solon (2003) and Ramos (2003) manage to incorporate both a random growth and a random walk term in their model. I am not able to do this, as the estimate of  $\sigma_{\beta}^2$  still is negative with this strategy.<sup>21</sup>

Next in Table 3 are the cohort specific parameters  $q_b$  in the permanent component. For identification, the value for the cohort born in 1934/35 is normalized to unity.<sup>22</sup> The estimates are larger for younger cohorts. A natural first interpretation of this is that younger cohorts have more heterogeneous skills. Consistent with this interpretation, Edin and Holmlund (1995) find increased dispersion in measured human capital characteristics from 1968 to 1991 for representative Swedish samples. Hence, with constant prices, wage inequality should have increased in Sweden.

Other interpretations of the  $q_b$  estimates are possible. It is likely that somewhat different skills are needed in the labor market in 1990 than in 1960. The skills needed for the 1990 labor market could then be more unequally distributed among all cohorts. The relative prices of these skills could also differ if they have a different marginal impact on firms' profit. With explanations along these lines, the factor loadings on the permanent component cannot stringently be interpreted as capturing changes in the price of skills, because the type of skills that it loads changes over time.

---

<sup>21</sup> The data that I use is richer than the data used by Baker and Solon (2003) and Ramos (2003). Hence, their ability to incorporate both specifications must be due to a different covariance structure of earnings in the U.K. and Canada than that for Sweden.

<sup>22</sup> I have also estimated a specification where only the initial shock at age 27 is allowed to differ between cohorts. The number of estimated parameters in this specification is the same as in the specification in Table 3, but the sum of squared residuals is always larger.

It could also be that younger cohorts consistently are hit by larger permanent shocks, possibly even though they do not have a larger dispersion of skills. This could be the case if the labor market has become ‘tougher’ over time. There has been a trend increase in Swedish youth unemployment (Edin and Holmlund, 1995), which at least does suggest that young individuals face a tougher labor market.

Yet another interpretation of the estimated  $q_b$  parameters is possible. Individuals included in the panel from older cohorts must have earnings above the basic amount in a later stage of life than younger cohorts. If individuals with lower skills drop out of the labor market over the life cycle, then the dispersion of skills in my sample will be higher for younger cohorts, even if this is not the case for the population as a whole. In the data, the sample size for a cohort generally decreases at older ages and those who drop out of the sample do generally have lower mean earnings. However, these effects mainly start around the age of 50. If sample selection were the sole explanation behind the results, one would not expect such large differences between cohorts who only are included in the sample while they are younger than 50 years old.

Next in Table 3 are the estimated factor loadings on the permanent component, denoted  $p_t$ . For identification, the value for 1960 is normalized to unity. The factor loadings generally decrease during the whole sample period but the fall is more rapid up until the early 1980s than afterwards. Holding the cohort parameters constant, this implies a trend decrease in permanent inequality between 1960 and 1990.

Next I turn to the estimates for the transitory component. First are the parameters of the ARMA-process. Both the AR- and the MA-parameter are larger for 1960-67 than for 1968-90. This may be due to the data differences, but it may also be due to real labor market differences. The values  $\hat{\rho}_{1960-67} = 0.81$  and  $\hat{\delta}_{1960-67} = -0.26$  imply that 55 percent of a transitory shock remains after 1 year, and that 22 percent remains after 5 years. The estimates

for 1968-90,  $\hat{\rho}_{1968-90} = 0.57$  and  $\hat{\delta}_{1968-90} = -0.26$ , imply that 31 percent remains after 1 year, and that 3 percent remains after 5 years. The estimates for 1968-90 are of the same magnitude as the corresponding estimate for Sweden and the period 1991-99 in Gustavsson (2004).<sup>23</sup> The magnitude is also similar to that obtained for the U.S. in Haider (2002).

The parameters that allow for age heteroskedasticity in the variance of the transitory innovation follow next. The negative  $\hat{\gamma}_1$  and the positive  $\hat{\gamma}_2$  imply a U-shaped age profile. The same result is found for Canada in Baker and Solon (2003) and for Sweden during 1991-99 in Gustavsson (2004). However, compared to these two studies, the implied differences between ages are small. The largest difference in the variances between two ages is here 12.5 percent, whereas the largest differences in Baker and Solon (2003) and Gustavsson (2004) is well-above 50 percent. At least part of this difference is likely to be explained by the differences in the applied earnings thresholds for the samples.

Next I turn to the estimated year-specific factor loadings on the innovation in the transitory process, denoted  $\lambda_t$ . Here, the factor loading for 1962 must be normalized to unity because the innovation variance in 1960 must be 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 estimates show that transitory shocks increased during the 1960s, steadily decreased from the early 1970s to the early 1980s, were fairly constant up to 1986, and strongly increased thereafter.

Last, the cohort specific initial transitory variances in Table 3, denoted  $\sigma_b^2$ , capture the accumulation of the transitory component up to the start of the sample period for each cohort. For the cohorts born 1912/13 to 1932/33 – who enter the sample in 1960 - these estimates reflect not only potential cohort heterogeneity but also age differences, as they enter the sample at different ages. It is hard to see any systematic differences in the estimates between

---

<sup>23</sup> Gustavsson (2004) only includes an AR-term.



these cohorts, the conclusion being that neither systematic age nor cohort effects are important.

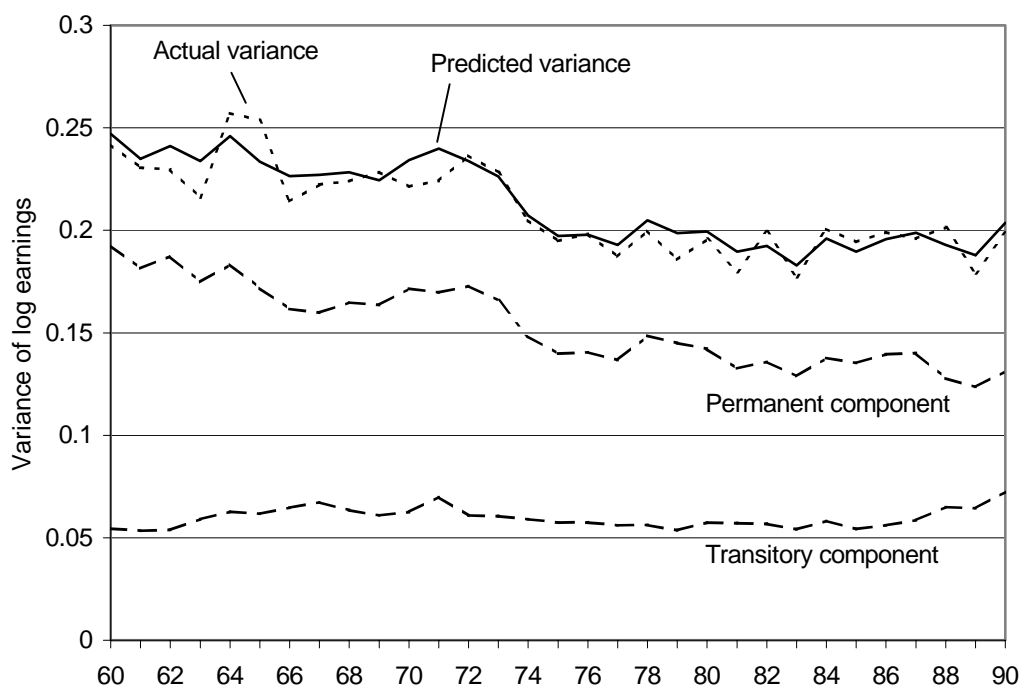
For the cohorts born 1934/35 and later, the cohort specific initial variances do not reflect age effects as all these cohorts enter the sample as of age 27. These cohorts instead enter the sample at different years and the estimates therefore capture time effects as well as cohort effects. It is possible to get a rough idea of the relative importance of cohorts and time effects for these cohorts by comparing changes in the estimated factor loadings on the transitory innovation to differences in the initial variances.<sup>24</sup> The factor loadings increase during the 1960s but this is also true for the initial variances for the cohorts that enter the sample during this decade, born between 1934/35 and 1940/41. It is hence hard to draw any conclusions for these cohorts. The rest of the cohorts, born 1942/43 to 1958/59, enter the sample between 1970 and 1986. The initial variances for these cohorts have roughly the same magnitude. The factor loadings, on the other hand, decrease during the 1970s and are constant during the first half of the 1980s. This indicates that among these cohorts, younger cohorts are more exposed to transitory earnings variability.

What do the estimates in Table 3 imply about changes in the permanent and transitory variance of earnings? To study this I follow Baker and Solon (2003) and use the full model to predict the permanent and transitory components each year for individuals 40 years old. This is the mid age in the sample and should correspond to individuals in the middle of their working careers. Figure 2 displays the predictions. In moving from year to year the factor loadings on each component change, and so do the cohort specific parameters.<sup>25</sup>

---

<sup>24</sup> Actually, one should look at the whole transitory component, except for the cohort effects, when doing the comparison. However, this series is very similar to the series for the factor loadings and is thus why these are used as an approximation.

<sup>25</sup> In fact, the cohort specific parameters change every two years. For example, in 1960/61 the cohort specific estimates for the cohort born 1920/21 are used as members of this two-year birth cohort are 40 years old these years. For 1962/63, the estimates for the cohort born 1922/23 are used.



**Figure 2: Decomposition of the variance of log earnings among men 40 years old.**

In Figure 2, the permanent component displays a downward trend during the whole sample period, but its decrease is much more rapid up until the early 1980s than afterwards. The transitory component steadily increases until the early 1970s and then decreases to be in a similar magnitude in the early 1980s as in the early 1960s. From the mid 1980s and on, it shows a clear upward trend. The evolution of the transitory component as a whole is smoother than that for the transitory shocks. This is because the shocks are dampened by the ARMA-process which smoothes them out because shocks also affect later earnings.

Figure 2 also contains the predicted and actual variances for 40-year olds.<sup>26</sup> The predicted variance misses some year-to-year movements in the actual variance during the

<sup>26</sup> The actual variances are chosen to be consistent with the cohort specific parameters used in the predictions; see footnote 25. Hence, for 1960, the actual variance for the two-year birth cohort born 1920/21 is a direct estimate of the variance for individuals defined to be 40 years old in that year. In 1961, I use the actual variance for the cohort born 1920/21, in 1962 I use the two-year cohort born 1922/23, and so forth. As the age of each two-year birth cohort is defined by the older of the cohorts, this means that the actual variances in the odd-numbered years correspond to individuals who in the estimations are defined to be 41 years old. These variances are adjusted downward with the in Table 3 estimated variance of the permanent shock as of age 41 and with the difference in the estimated transitory variance between the age of 41 and 40.

1960s but captures the main pattern during the decade. After the 1960s, the predicted variance quite closely captures the movements in the actual variance.

I have also performed decompositions for 30 and 50 year-olds (available on request). Changes during the 1960s and the 1970s are similar to those for 40-year olds, but some differences stand out during the 1980s. For 30 year-olds, both the permanent component and the predicted/actual variance increase noticeably after 1983. For 50 year-olds, the permanent component is constant during the second half of the 1980s, and as a result, the predicted/actual variance increases slightly.

It is important to remember that this study, like those for the U.S., only investigates earnings inequality among men. The substantial increase in female labor force participation in Sweden during the studied period may have affected permanent earnings inequality among men. For example, Topel (1994) finds for the U.S. labor market that high skilled women are better substitutes for low-skilled men than are low-skilled women. The inflow of high-skilled women into the labor force could then increase wage differentials between low and high-skilled men. On the other hand, Juhn and Kim (1999) find little evidence that women substitute for men or that they have contributed to rising wage inequality among men, at least in the U.S. labor market. I am not aware of any Swedish study that looks at the effect of rising female labor supply on male wages.

The increase in the transitory variance of earnings during the second half of the 1980s has an interesting link to wages. According to Edin and Holmlund (1995), a majority of the increase in wage dispersion during the second half of the 1980s was due to increased within-group dispersion. This could be due to increased returns to (for the econometrician) unobserved skills, but it could also be due to increased earnings instability. The results in this paper support the latter explanation.

When the permanent variance has decreased, the transitory variance has increased or held constant. Hence, the relative importance of the transitory variance has increased. This reflects decreases in the auto-correlations of earnings, which in turn implies that individuals more often shift positions in the earnings distribution in 1990 than in 1960. A study that directly estimates the evolution of income mobility in Sweden between 1951 and 1989 is Björklund (1993). My results are fully consistent with his results.<sup>27</sup>

An interesting question is whether there is any trend in the transitory component as measured over the whole 31 years. It is also important to investigate whether the unemployment and sickness absence rate affect the estimates, as sickness benefits and unemployment benefits are not included in the earnings measure. To do this, I follow Baker and Solon (2003) and apply least squares to estimate time-series regressions of the transitory and permanent components between 1960 and 1990 for 40 year-olds using a linear time trend, the unemployment rate, and sickness absence as regressors. The regression results are reported in Table 4. For the transitory component, none of the explanatory variables are significantly different from zero. For the permanent component there is a strongly significant negative trend, but unemployment and sickness absence have no significant effect.<sup>28</sup> The insignificant effect of the unemployment rate may be due to its small variation; the highest unemployment rate in a single year between 1960 and 1990 was 3.7 percent in 1983.

The lack of a stable connection between the two components and changes in unemployment and sickness absence has important implications for the interpretation behind their changes. Estimated changes during part of the sample period, such as during the 1960s,

---

<sup>27</sup> The results in Björklund (1993) are, however, not fully comparable because the income measure used is the sum of labor earnings, capital income, and realization of capital gains, and from 1974 also taxable benefits such as parental leave payment, unemployment benefits, and pensions.

<sup>28</sup> The same results apply to 30 and 50 year-olds. The measure of unemployment and sickness absence used contains measurement errors due to several changes in the definitions during the sample period. Unfortunately, no better data is available.

**Table 4: The impact from trend, unemployment, and sickness absence on the permanent and transitory components 1960-90<sup>a</sup>**

| Dependent variable   | Linear trend         | Unemployment rate <sup>b</sup> | Sickness absence <sup>c</sup> | Constant           | R <sup>2</sup> | Durbin-Watson |
|----------------------|----------------------|--------------------------------|-------------------------------|--------------------|----------------|---------------|
| Transitory component | 5.4E-05<br>(2.4E-04) | -0.2207<br>(0.2338)            | 1.9E-4<br>(5.3E-4)            | 0.0597<br>(0.0104) | 0.08           | 0.69          |
| Permanent component  | -0.0019<br>(3.1E-4)  | 0.0451<br>(0.3722)             | -0.0019<br>(3.1E-4)           | 0.1941<br>(0.0119) | 0.86           | 1.02          |

<sup>a</sup> OLS estimates. Dependent variables are calculated for males 40-year old based on the estimates in Table 3. Newey-West robust standard errors are in parantheses

<sup>b</sup> Source 1963-90: Labor force surveys (AKU). Source 1960-62: Statistics Sweden, Statistical Yearbooks. The measure 1960-62 has been re-scaled to better match the measure during later years.

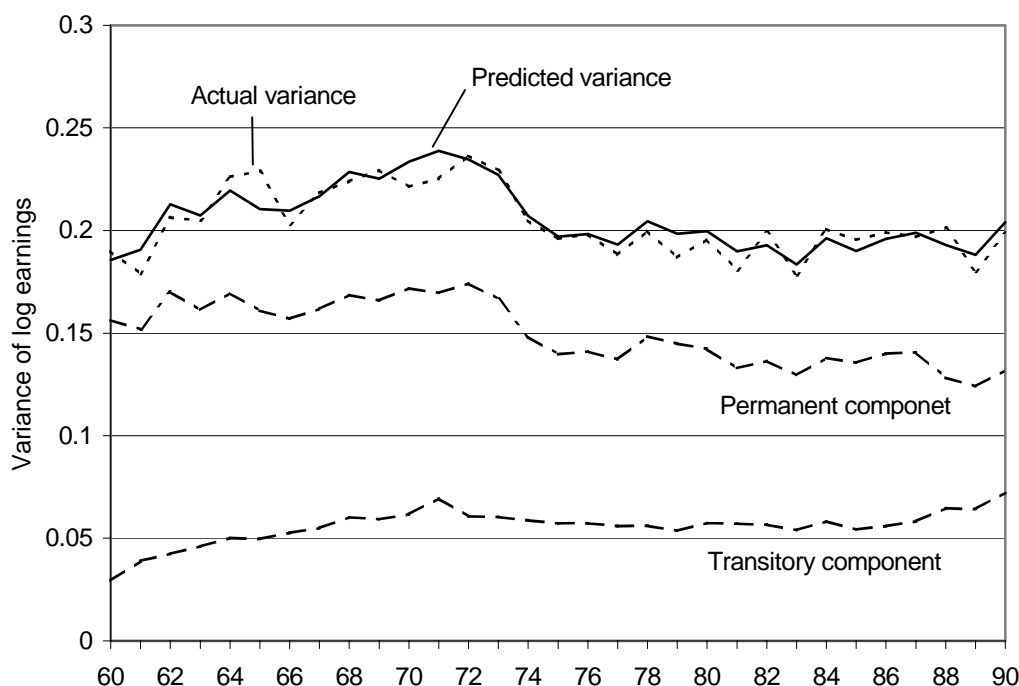
<sup>c</sup> The variable measures sickness absence ("sjuktalet") for males 16-64 years old. The data used and further information can be found at [www.RFV.se](http://www.RFV.se).

are likely to be due to real changes in the functioning of the labor market rather than to short term fluctuations in hours worked linked to unemployment or sickness absence.

#### 4.2 Sensitivity analysis

This sub-section contains six sensitivity analyses. First, I discuss how the results change when the Pareto distribution is applied to the top-coded earnings for 1960-67. Second, I compare the results obtained with artificially top-coded earnings for 1968-75 with the results from correct earnings for the same period. Third, I look at whether the results change when separate estimations are performed for the 1960-67 and 1968-90 periods. Fourth, I report results based on a balanced revolving panel. Fifth and sixth, I discuss how the estimates change when a lower earnings threshold and an earnings threshold adjusted for changes in real wages are used in the sample construction.

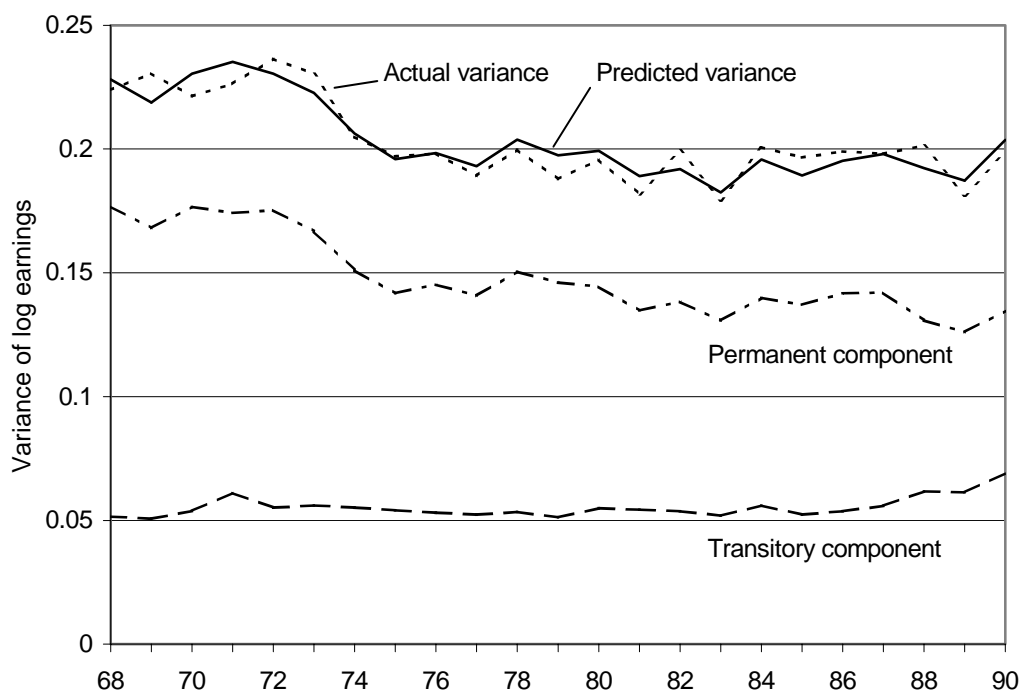
The top-coded data for the 1960-67 period used in the estimations reported in Table 3 were replaced with their implied mean value based on the lognormal distribution. An alternative choice, as discussed above, is the Pareto distribution. The results in Appendix B indicate that auto-correlations are more correctly estimated when this distribution is applied. This is important since changes in the *relative* size of the permanent and transitory components are identified by changes in auto-correlations. However, the results also show that the Pareto distribution tends to overestimate the variances when the percentage top-coded



**Figure 3: Decomposition of the variance of log earnings among men 40 years old, where top-coded earnings have been mean-approximated based on the Pareto distribution.**

is around the level present for the 1964-67 period. Based on mean-approximated data for which the Pareto distribution has been used, but without re-scaling of the 1960-67 earnings because of the overestimation problem, Figure 3 reports the decomposition of the variance of log earnings for 40 year old males based on the estimation of the model in equations (5)-(8). The predicted/actual variance increases during 1960-67, whilst actual evidence instead indicates that earnings inequality was constant between those years. However, what is of more interest here is the evolution of the relative size of the transitory component and it is clear that it increased between 1960 and 1967. Given constant earnings inequality between 1960 and 1967, this is evidence that the transitory component indeed did increase during this period and that the permanent component indeed did decrease.

Appendix C contains results for the 1968-90 period based on data where earnings for 1968-75 have been artificially top-coded and exposed to sample selection in such a way as to mimic the 1960-67 data. The use of mean-approximated data based on the lognormal distribution and re-scaling is found to produce estimates of changes in the transitory and



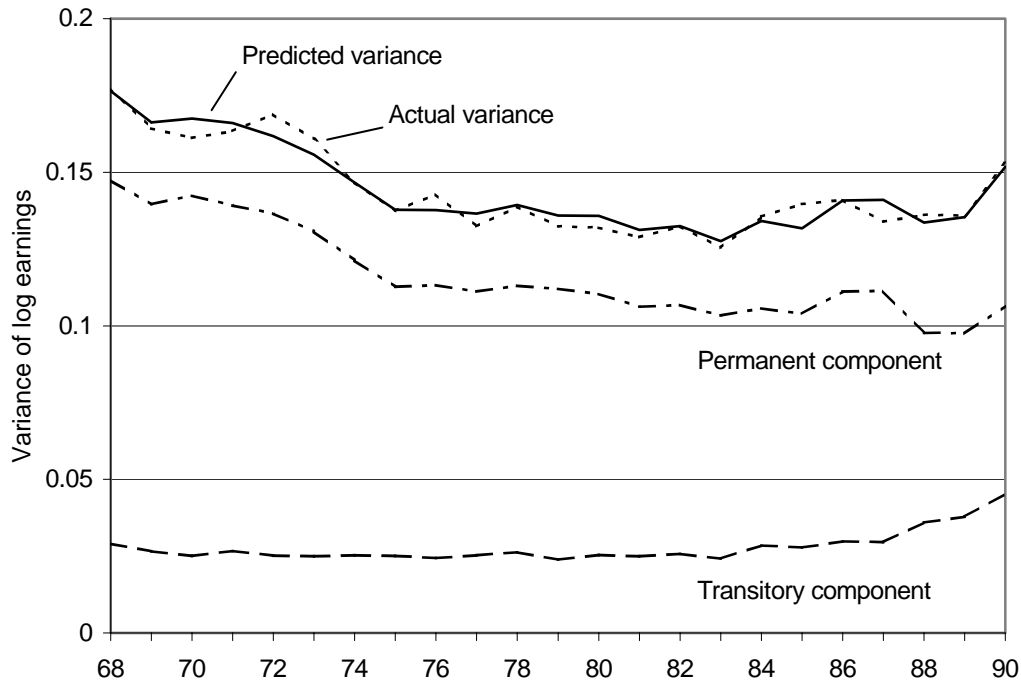
**Figure 4: Decomposition of the variance of log earnings among men 40 years old, based on a 1968-90 panel.**

permanent components that are reasonably correct. However, the relative size of the transitory component tends to be overestimated for the years that top-coded data are used.

Could potential biases during 1960-67 have carried over to the results for 1968-90 in Table 3? To test this, I have estimated the model in equations (5)-(8) using only data for 1968-90.<sup>29</sup> Figure 4 presents the predicted permanent and transitory components for 40 year-olds. The relative size of the transitory component is smaller during 1968-73; the 1960-67 data has hence had some effect on the results for later years. However, both components show the same pattern over time as before and no main conclusions are altered.

Another issue is the use of unbalanced panels. With this design, slightly different samples are used over time when a cohort's variances and auto-covariances are estimated. To make sure that changes in sample composition are not driving the main results, I have also performed the 1968-90 estimations with a revolving balanced panel. The parameter estimates

<sup>29</sup> The same sample selection procedure is applied as before, for example, only two-year birth cohorts aged 26-53 for at least 6 years between 1968 and 1990 are included.



**Figure 5: Decomposition of the variance of log earnings among men 40 years old, based on a 1968-90 revolving balanced panel.**

based on this panel tells the same main story concerning age and cohort effects as the estimates from the unbalanced panels. Figure 5 graphs the decomposition for 40 years old males. The variances are smaller, especially the transitory variance. The transitory variance is also more stable up to 1983 but from that year onwards displays a much stronger increase than that produced by the unbalanced panels. The main difference for the permanent component is that its rapid fall ends in 1975, earlier than is the case for the unbalanced panels.

I have also estimated models based only on data for the years 1960-67, with and without the re-scaling of the mean-approximated top-coded earnings. These models have to be much simpler because with data only for 1960-67 there are fewer elements in each cohort specific covariance matrix.<sup>30</sup> However, for the models that I am able to fit, both the results for an

<sup>30</sup> To break the direct connection between age and time in the sample, all two-year birth cohorts between the ages 27-52 for at least 3 years 1960-67 are included, which makes it hard to estimate cohort effects. The upper age in this sample has to be 52 because the oldest cohort available in the PIR is born in 1910, and these individuals are 50 years old in 1960 and hence 52 years old after 3 years in the sample. The inclusion of the additional cohort born 1910 induces more sample selection and the lack of cohort specific estimates in the models that I am able to estimate means that I cannot control for this.



unbalanced and a balanced panel produce the same conclusions as before for this period (available on request).<sup>31</sup>

In the samples used so far, individuals are only included in a given year if they have earnings above the basic amount in that year, where the basic amount in 1960 corresponds to 41,143 SEK in 2002. To see how a lower earnings threshold affects the estimates, I have also used a 1968-90 sample with a threshold corresponding to 27,595 SEK in 2002; this is the lowest threshold I can use because individuals with lower earnings were not obligated to fill a tax report.<sup>32</sup> The results based on this sample are similar to those for the sample with the higher threshold; see Figure D1 in Appendix D for the decomposition into permanent and transitory components. One difference is that the transitory component does not show any trend during the 1970s. The relative size of the transitory component and the age variation in the variance of the transitory innovation is also slightly larger.

Another issue is the use of a constant earnings threshold in the sample. Real wages increased noticeably from 1960 to the mid 1970s to be roughly constant thereafter. Individuals therefore generally have to work fewer hours to be included in the 1990 sample than in the 1960 sample. However, that revolving balanced panels, in which an individual must have earnings each year above the basic amount to be included in the panel, give similar results as unbalanced panels is a good indication that the constant threshold is not driving the main results. I have also experimented with an earnings threshold that is adjusted for changes in real wages. However, the real wage index that I am able to create only applies to workers in manufacturing, and the definition of wages used changes several times. It is also hard to draw any conclusions for the period 1960-67 with such data because the percentage top-coded becomes higher as non-censored observations must be discarded. The result of this is that the mean-approximation method performs less well. Because there are no top-coded observations

---

<sup>31</sup>No estimations with re-scaled data have been conducted for the balanced panel, as it is far from clear how the requirement of consecutive years of earnings should be incorporated into the 1968 data.

<sup>32</sup>This is the threshold for filling a tax report in 1971; other years have somewhat lower thresholds.

between 1968 and 1990, it is more straightforward to draw conclusions for this period. None of the conclusions for this period change with data in which the earnings threshold each year corresponds to the basic amount in 1960 adjusted for changes in real wages; see Figure D2 in Appendix D.

Overall, the results presented in this sub-section support the conclusion that the transitory variance of earnings increased from 1960 to the early 1970s, slightly decreased or held constant until around 1980, then increased during the second half of the 1980s. The permanent variance decreased strongly until the early 1980s, after which it declined at a slower pace.

## **5. Discussion and a comparison with the U.S. evidence**

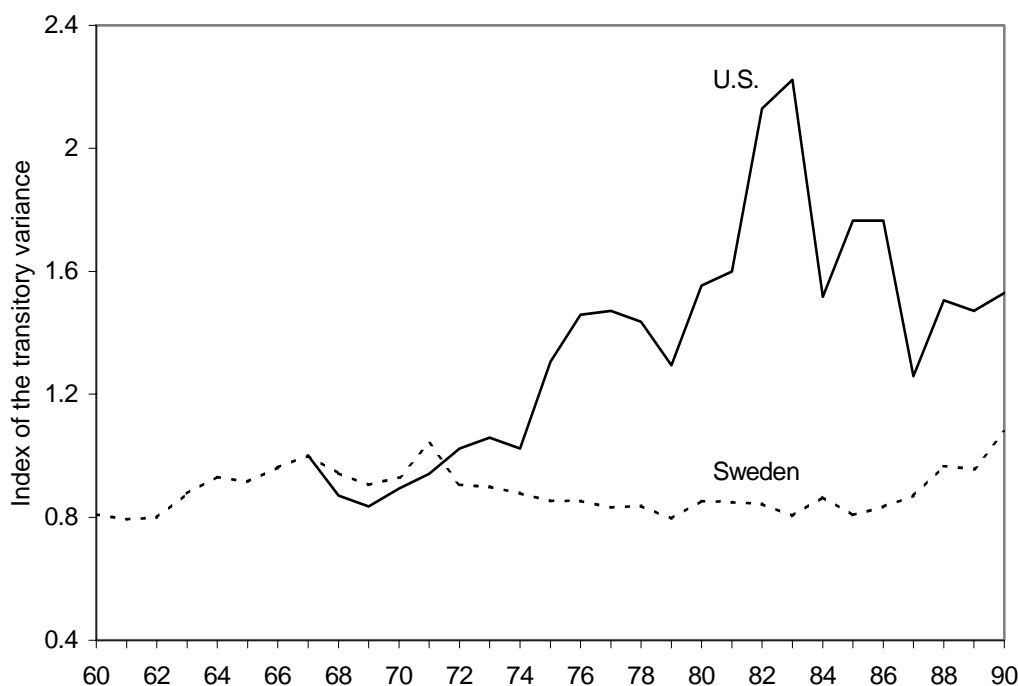
Differences in the evolution of the returns to skills, for example to education, between Sweden and the U.S. are commonly known and have been the subject of much research (see e.g. Freeman and Katz, 1995). Because changes in the returns to skills are closely related to changes in permanent earnings inequality, this section mainly compares changes in the transitory variance of earnings.<sup>33</sup>

Figure 6 displays the evolution of the transitory variance of log earnings, or earnings instability, in the U.S. between 1967 and 1990 for white males based on the results in Haider (2001) along with similar estimates for Sweden between 1960 and 1990 based on the results in Table 3. The value for 1967 is normalized to unity in both series.

The difference between Sweden and the U.S. during the 1970s is striking. According to Haider's estimates, U.S. earnings instability strongly increased from the beginning of the 1970s to the early 1980s, and the increase was clearly much stronger than anything

---

<sup>33</sup> For a comparison of earnings mobility between the U.S. and Sweden, see Aaberge *et al* (2002). Their definition of earnings differs in several respects to that used here.



**Figure 6: The evolution of the transitory variance of log earnings in the U.S. between 1967 and 1990 and in Sweden between 1960 and 1990.**

Source: The numbers for the U.S. are based on Haider (2001), Table 5, column 4, p. 821. The numbers for Sweden are based on Table 2.

Note: The value for 1967 is normalized to unity in both series. The numbers for Sweden are for 40 year-olds and cohort heterogeneity is included in the predictions. Haider (2001) does not allow for age variation or cohort heterogeneity in the transitory component of earnings.

experienced in Sweden between 1960 and 1990.<sup>34</sup> It is clear that either something happened in the U.S. labor market that did not occur in Sweden, or it had a much smaller impact in Sweden because of differences in the functioning of the labor markets. The experiences of the two countries also diverged during the 1980s, as instability decreased slightly in the U.S. but increased in Sweden from the mid 1980s. Haider's results also show that permanent earnings inequality in the U.S. rose sharply from 1979 through to the second half of the 1980s; see Figure D3 in Appendix D for a comparison to Sweden.

<sup>34</sup> Of course, part of this can be due to a much lower earnings threshold in Haider (2001), but the estimates that I have performed with a lower threshold did not predict more powerful trends in the transitory variance; see Figure D1 in Appendix D. Based on the results in Haider (2001), I have estimated a time-series regression of the transitory component (normalized to unity in 1967) on a constant, a linear time trend, and the unemployment rate. A one-percentage point increase in the unemployment rate is significantly associated with a 0.14 increase in the normalized transitory component. Part of the higher earnings instability in the U.S. in the first half of the 1980s compared to the early 1970s is therefore likely to be due to a higher unemployment rate, but this can only explain part of the strong increase.

Haider (2001) also reports results for average wages during a year, i.e. annual earnings divided by hours worked. The estimate of the transitory variance of wages displays a noticeable decrease in the few years between 1967 and 1969, a trend increase between 1969 and 1983, and is quite constant thereafter. One notable difference from annual earnings is that most of the increase in the transitory variance occurs in the first half of the 1970s. The results for the permanent component are similar for average wages and annual earnings.

Moffitt and Gottschalk (2002) use annual earnings and the same methodology as Haider (2001) and reach similar results.<sup>35</sup> For earnings instability, one difference is that they find that most of the increase occurred from the late 1970s until around 1985, whereas Haider (2001) find that the increase started in the mid 1970s and ended around 1983. As both studies use the PSID, the differences are likely to be due to different sample constructions.<sup>36</sup> Gottschalk and Moffitt (1994) use a different methodology and find that annual earnings instability was on average 42 percent higher during 1979-87 than during 1970-78. They also find that wage instability substantially increased, being on average 35 percent higher during 1979-87 than during 1970-78. Gittleman and Joyce (1996) use yet another methodology and their results are consistent with an increase in annual earnings instability from the late 1970s and during the first half of the 1980s.

Overall, annual earnings instability strongly increased from around the mid 1970s to the mid 1980s in the U.S. while it decreased or was constant in Sweden during this period. What could explain the difference? Gottschalk and Moffitt (1994) investigate several factors that might explain the U.S. increase in earnings instability. They find that factors like de-unionization and changing industrial composition explain only about 12 percent of the increase. Likewise, increases in job turnover, self-employment, and part-time work cannot

---

<sup>35</sup> I am not able to graph the evolution of earnings instability as estimated in Moffitt and Gottschalk (2002) because they do not report their parameter estimates.

<sup>36</sup> Haider (2001) studies white males during 1967-91 aged 25-60 using a revolving balanced panel but Moffitt and Gottschalk (2002) study white and non-white males during 1969-96 aged 20-59 using a fully unbalanced panel.

account for the majority of the increase in earnings instability. As these factors do not appear to explain much, I will not compare each of these between the U.S. and Sweden. Instead, I will first discuss two broad hypotheses: skill-biased technological change and increased globalization. These are often discussed in connection with permanent earnings inequality, but also hold some promise in explaining increased earnings instability.<sup>37</sup> I then turn to differences in labor market institutions.

Several studies (Autor *et al.*, 1998; Machin and Van Reenen, 1998; Bresnahan *et al.*, 2002) have found evidence consistent with the hypothesis that from the mid 1970s to the mid 1990s, the U.S. has undergone a period of accelerating skill-biased technological change associated with increased R&D intensity, the spread of computer-based technologies, and new organizational practices, and that these developments could explain part of the increase in the U.S. permanent earnings inequality. Note, however, that not all evidence is consistent with this hypothesis (see Card and DiNardo, 2002).

Besides affecting permanent earnings inequality, Katz and Autor (1999) raise the possibility that accelerating skill-biased change could also, at least in a transition period, generate increased earnings instability since firms are likely to have much initial uncertainty concerning the abilities of individual workers to perform new tasks and adapt to new organizational environments. Hence, if accelerating skill-biased technological change started earlier in the U.S. than in Sweden, this could potentially explain the diverging pattern for earnings instability. However, Machin and Van Reenen (1998) report that R&D intensity in manufacturing grew at least as much in Sweden as in the U.S. during the period 1973-89. The results in Hansson (1997) are also consistent with the presence of this type of skill-biased technological change in Sweden during the 1970s. Differences in skill-biased technological

---

<sup>37</sup> These two hypotheses focus on changes in labor demand. It is possible that changes in labor supply can affect earnings instability, but I am not aware of any potential mechanisms through which this might occur.

change, *given the existing research*, therefore do not appear to be the main explanation for the diverging evolution of earnings instability in the U.S. and Sweden.

Katz and Autor (1999) also raise the possibility that increased globalization may increase earnings instability by making labor demand curves more elastic so that shocks to product market prices have a larger impact on wages. However, as far as increases and changes in the pattern of international trade are concerned, Sweden and the U.S. do not differ greatly during the 1970s and 1980s. Moreover, evidence suggests that changes in international trade might have had a minor impact on the functioning of the Swedish and U.S. labor markets; see Hansson (1997, 2000) for Sweden and the discussion and references in Katz and Autor (1999) for the U.S. experience.

To conclude, the increase in U.S. earnings instability from the mid 1970s to the mid 1980s appears to coincide with a period of accelerating skill-biased technological change. This period has also seen changed patterns in international trade. The same is also true for Sweden, but earnings instability did not increase during this period. One possible explanation for this is, of course, that these factors do not affect earnings instability. Another possible explanation is that Sweden's labor market institutions, in particular, centralized wage setting and egalitarian wage policy, imposed constraints on earnings instability.

The solidarity wage policy became more egalitarian from the late 1960s, evolving from a goal of "equal pay for equal work" to just "equal pay" (Edin and Topel, 1997; Hibbs and Locking, 2000). Given this and the comparison to the U.S., it is striking that earnings instability increased in Sweden during the 1960s and during the second half of the 1980s, two periods when the strongly egalitarian version of the solidarity wage policy had less bearing on

the labor market. In contrast, earnings instability decreased or was constant during the policy's heyday in the 1970s and the early 1980s.<sup>38</sup>

At this point it is informative to look at results for earnings instability that are available for other countries. Dickens (2000) contains results for the U.K. between 1975 and 1995 and Baker and Solon (2003) present results for Canada between 1976 and 1992. In both countries, as in the U.S., earnings instability increased noticeably during the first half of the 1980s - a period during which earnings instability was constant in Sweden. In terms of wage setting institutions, these countries have more in common with the U.S. than with Sweden. Cappellari (2004) contains results for earnings instability in Italy between 1979 and 1995 which indicate that the transitory variance of earnings decreased in Italy during the 1980s. In terms of wage setting institution, Italy is more like Sweden than the U.S. In fact, Italy has a history of wage compression that strongly resembles that of Sweden (see Erickson and Ichino, 1995).<sup>39</sup>

For Sweden, there are several mechanisms through which centralized solidarity bargaining may have reduced earnings instability. In terms of skill-biased technological change, it must have been harder for an employer in Sweden than in the U.S. to adjust wages to the inflow of new information about workers' abilities, at least when it comes to wage cuts. For changed patterns in international trade, Edin and Topel (1997) argue that the centralized solidarity bargaining to some extent pushed market forces out of play. Thus it is possible that more elastic labor demand curves had less impact on the Swedish labor market. Other mechanisms are also possible. The solidarity wage policy meant that workers' wages were quite independent of firm or industry profit, and therefore would fluctuate less than in the

---

<sup>38</sup> Gustavsson (2004) finds that transitory shocks to earnings in Sweden continued to increase during the first years of the 1990s. However, the deep Swedish recession at the time makes it hard to say if anything else but this caused the increase.

<sup>39</sup> Italy experienced a substantial union driven compression of wage differentials during the 1970s. Like Sweden, wage compression came to a stop around 1983. But unlike Sweden, there is no clear-cut evidence of increased cross-sectional wage dispersion in Italy during the 1980s (Erickson and Ichino, 1995).

unconstrained U.S. labor market.<sup>40</sup> Because of the overall wage compression in Sweden, workers who frequently changed jobs, voluntarily or involuntarily, would receive more equal wages at each job than the corresponding worker in the U.S. labor market, with less earnings instability as a result.

It could, however, be argued that the increase in earnings instability during the 1960s was by itself a consequence of the solidarity wage policy. Results in Holmlund (1976) indicate that the Swedish labor market was, to some extent, turbulent during the second half of the 1960s. For example, there were generally more labor layoffs than during the first half of the 1970s. Edin and Topel (1997) also report that regional migration rates were unusually high from the mid 1960s to the early 1970s and argue that the solidarity wage policy had allocative effects on the labor market. There is, however, no direct evidence that the solidarity wage policy caused the labor market to be turbulent. Another problem with the wage policy based explanation for the 1960s is that the original version of the solidarity wage policy, which aimed for “equal pay for equal work”, was prevalent up until the late 1960s. If the “equal pay for equal work” policy increased earnings instability through labor market turbulence, one would expect the movement to just “equal pay” at the late 1960s to at least keep earnings instability at a similarly high level. Instead, earnings instability began to decrease after 1971, even though the egalitarian version of the solidarity wage policy markedly compressed wages at least until the mid 1970s. It is hence uncertain what the underlying causes for the increase during the 1960s were. The lack of international comparisons also makes it harder to draw conclusions.

---

<sup>40</sup> A large fraction of the total wage increases actually occurred through wage drift at the firm level, i.e. wage increases beyond that stated in the central agreement. Some studies (e.g. Flanagan, 1990; Jackman, 1990) have argued that this eroded the impact of centralized bargaining. This would mean that wages were directly affected by industry and firm profit. However, Hibbs and Locking (1996) show that wage drift was accurately predicted and fully incorporated into central wage agreements between the blue collar union LO and the employer organization (SAF), and hence incorporated into the solidarity wage policy.



One could also argue that earnings instability in Sweden decreased during the 1970s due to strong employment growth in the Swedish public sector at the time. Such a claim is supported by the results in Cappellari (2002), who finds for Italy that the transitory variance of earnings is smaller in the public sector than in the private sector. However, Rosen (1997) show that women accounted for all of the employment growth in the Swedish public sector. Hence, this employment growth should not directly affect my estimates, as I only study men.

To conclude, there are strong reasons to believe that centralized solidarity bargaining imposed constraints on earnings instability in Sweden during most of the 1970s and the early 1980s.

## **6. Concluding remarks**

Using a large individual longitudinal database, I decompose the cross-sectional variance of male annual earnings in Sweden between 1960 and 1990 into permanent and transitory components. The transitory variance has increased from 1960 until the early 1970s, decreased slightly until the late 1970s, and then risen again during the second half of the 1980s. The permanent variance displays a downward trend during the whole sample period, with the rate of decline more rapid up until the early 1980s than afterwards.

A comparison of the evolution of the transitory variance in Sweden with evidence for the U.S. reveals sharp differences. In the U.S., earnings instability increased sharply from the mid 1970s to the mid 1980s, with the increase more rapid than anything experienced in Sweden between 1960 and 1990. Gottschalk and Moffitt (1994) find that factors such as changed job turnover and industry composition do not explain much of the U.S. increase. Thus, differences in these particular factors appear unlikely to explain much of the difference with Sweden. The large differences observed between Sweden and the U.S. are also unlikely

to be explained by differences in skill biased technological change or changes in the pattern of international trade.

The comparison between Sweden and the U.S. opens up for the possibility that labor market institutions may be significantly responsible for differing evolution of earnings instability. In particular, it is likely that centralized solidarity bargaining in Sweden imposed constraints on earnings instability during the 1970s and early 1980s. The results in Edin and Zetterberg (1992) and Edin and Topel (1997) suggest that a large share of the differences in permanent earnings inequality between Sweden and the U.S. can be explained by differences in labor market institutions. That labor market institutions affect the evolution of transitory earnings inequality does therefore not appear to be a far-fetched hypothesis.

Unfortunately, it is harder to make concrete statements about the underlying causes of Swedish earnings instability during the 1960s, as there are no international studies to place it in context. There is evidence that the Swedish labor market, to some extent, was turbulent during the second half of the 1960s, and this could be tied to the increase in earnings instability during this period.

Estimated changes in earnings instability can be given a welfare interpretation. As optimizing agents wish to smooth consumption over transitory earnings fluctuations, increased fluctuations are likely to decrease individual welfare. Only if one assumes that individuals have perfect foresight and can borrow and lend freely will earnings instability be welfare neutral. With more realistic assumptions, such as the presence of uncertainty and the existence of liquidity constraints, increases in earnings instability decrease individual welfare.

There are several reasons for believing that the strong egalitarian wage policy pursued by Swedish unions has distorted the Swedish labor market and been harmful to the Swedish economy; see Lindbeck *et al* (1993) and Edin and Topel (1997). However, if the pursued wage policy has indeed imposed constraints on earnings instability, then it must be viewed as

yielding a positive contribution to welfare by permitting less year-to-year fluctuation in individuals' consumption.

## References

- Aaberge, R, A Björklund, M Jäntti, M Palme, P Pedersen, N Smith and T Wennemo (2002), "Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States", *Review of Income and Wealth*, Vol. 48, pp. 443-469.
- Altonji, J and L Segal (1996), "Small-Sample Bias in GMM Estimation of Covariance Structures", *Journal of Business and Economic Statistics*, Vol. 14, pp. 353-366.
- Autor, D, L Katz and A Krueger (1998), "Computing Inequality: Have Computers Changed the Labor Market?", *Quarterly Journal of Economics*, Vol. 113, pp. 1169-1213.
- Baker, M (1997), "Growth Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings", *Journal of Labor Economics*, Vol. 15, pp. 537-579.
- Baker, M and G Solon (2003), "Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Tax Records", *Journal of Labor Economics*, Vol. 21, pp. 289-321.
- Björklund, A (1993), "A Comparison Between Actual Distribution of Annual and Lifetime Income: Sweden 1951-89", *Review of Income and Wealth*, Vol. 39, pp. 377-386.
- Bresnahan, T, E Brynjolfsson and L Hitt (2002), "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence", *Quarterly Journal of Economics*, Vol. 117, pp. 339-376.
- Cappellari, L (2002), "Earnings Dynamics and Uncertainty in Italy: How do They Differ Between the Private and Public Sectors?", *Labour Economics*, Vol. 9, pp. 477-496.
- Cappellari, L (2004), "The Dynamics and Inequality of Italian Men's Earnings: Long-term Changes or Transitory Fluctuations?", forthcoming in *Journal of Human Resources*.
- Card, D and J DiNardo (2002), "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles", *Journal of Labor Economics*, Vol. 20, pp. 733-783.
- Chamberlain, G (1984), "Panel Data", in Z Griliches and 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.
- Cowell, F (2000), "Measurement of Inequality", in A Atkinson and F Bourguignon (eds.), *Handbook of Income Distribution*, Vol. 1, Elsevier Science B.V.
- Dickens, R (2000), "The Evolution of Individual Male Earnings in Great Britain: 1975-95", *Economic Journal*, Vol. 110, pp. 27-49.
- Edin, P-A and P Fredriksson (2000), "LINDA – Longitudinal Individual Data for Sweden" Working Paper 2000:19, Department of Economics, Uppsala University.

- Edin, P-A and B Holmlund (1995), "The Swedish Wage Structure: The Rise and Fall of Solidarity Wage Policy?", in R Freeman and L Katz (eds.), *Differences and Changes in Wage Structures*, University of Chicago Press.
- Edin, P-A and R Topel (1997), "Wage Policy and Restructuring: The Swedish Labor Market Since 1960", in R Freeman, R Topel and B Swedenborg (eds.), *The Welfare State in Transition*, The University of Chicago Press.
- Edin, P-A and J Zetterberg (1992), "Interindustry Wage Differentials: Evidence from Sweden and a Comparison with the United States", *American Economic Review*, Vol. 82, pp. 1341-1349.
- Erickson, C and A Ichino (1995), "Wage Differentials in Italy: Market Forces, Institutions, and Inflation", in R Freeman and L Katz (eds.), *Differences and Changes in Wage Structures*, University of Chicago Press.
- Fichtenbaum, R and H Shahidi (1988), "Truncation Bias and the Measurement of Income Inequality", *Journal of Business and Economic Statistics*, Vol. 6, pp.335-337.
- Fitzgerald, J, P Gottschalk and R Moffitt (1998), "An Analysis of Sample Attrition in Panel Data: The Michigan Study of Income Dynamics", *Journal of Human Resources*, Vol. 33, pp. 251-299.
- Flanagan, R (1990), "Centralized and Decentralized Pay Determination in the Nordic Countries", in L Calmfors (ed.), *Wage Formation and Macroeconomic Policy in the Nordic Countries*, SNS.
- Freeman, R and L Katz (1995), *Differences and Changes in Wage Structures*, University of Chicago Press.
- Gittleman, M and M Joyce (1996), "Earnings Mobility and Long-Run Inequality: An Analysis Using Matched CPS Data", *Industrial Relations*, Vol. 35, pp. 180-196.
- Gottschalk, P and R Moffitt (1994), "The Growth of Earnings Instability in the U.S. Labor Market", *Brookings Papers on Economic Activity*, Vol. 1994 (2), pp. 217-272.
- Gustavsson, M (2004), "Changes in Permanent and Transitory Earnings Inequality in Sweden: 1991-1999", in M Gustavsson, *Empirical Essays on Earnings Inequality*, Ph.D. dissertation, Uppsala University.
- 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.
- Hansson, P (1997), "Trade, Technology and Changes in Employment of Skilled Labour in Swedish Manufacturing" in J Fagerberg, P Hansson, L Lundberg and A Melchior (eds.) *Technology and International Trade*, Edward Elgar Publishing Ltd.
- Hansson, P (2000), "Relative Demand for Skills in Swedish Manufacturing: Technology or Trade?", *Review of International Economics*, Vol. 8, pp. 533-555.

- Hibbs, D (1990), “Wage Dispersion and Trade Union Action in Sweden”, in I Persson (ed.), *Generating Equality in the Welfare State – The Swedish Experience*, Norwegian University Press.
- Hibbs, D and H Locking (1996), “Wage Compression, Wage Drift and Wage Inflation in Sweden”, *Labour Economics*, Vol. 3, pp. 109-141.
- Hibbs, D and H Locking (2000), “Wage Dispersion and Productive Efficiency: Evidence for Sweden”, *Journal of Labor Economics*, Vol. 18, pp. 755-782.
- Hofer, H and A Weber (2002), “Wage Mobility in Austria 1986-1996”, *Labour Economics*, Vol. 9, pp. 563-577.
- Holmlund, B (1976), *Arbetslöshet och lönebildning*, Ph.D. dissertation, Umeå University.
- Jackman, R (1990), “Wage Formation in the Nordic Countries from an International Perspective”, in L Calmfors (ed.), *Wage Formation and Macroeconomic Policy in the Nordic Countries*, SNS
- Jovanovic, B (1979), “Job Matching and the Theory of Turnover”, *Journal of Political Economy*, Vol. 87, pp. 972-990.
- Juhn, C and D Kim (1999), “The Effect of Rising Female Labor Supply on Male Wages”, *Journal of Labor Economics*, Vol. 17, pp. 23-48.
- Kakwani, N (1980), *Income Inequality and Poverty*, Oxford University Press.
- Katz, L and D Autor (1999), “Changes in the Wage Structure and Earnings Inequality” in O Ashenfelter and D Card (eds), *Handbook of Labor Economics*, Vol. 3A, Elsevier Science B.V.
- Leamer, E (1983) “Let’s Take the Con Out of Econometrics”, *American Economic Review*, Vol. 75, pp. 31-43.
- Linbeck, A, P Molander, T Persson, O Peterson, A Sandmo, B Swedenborg and N Thygesen (1993), “Options for Economic and Political Reform in Sweden”, *Economic Policy*, Vol. 8, pp. 219-263.
- Lydall, H (1968), *The Structure of Earnings*, Clarendon Press.
- Machin, S and J V Reenen (1998), “Technology and Changes in Skill Structure: Evidence from Seven OECD Countries”, *Quarterly Journal of Economics*, Vol. 113, pp. 1215-1244.
- 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 and P Gottschalk (1995), “Trends in the Covariance Structure of Earnings in the U.S.: 1969-1987”, mimeo, Johns Hopkins University.

- Moffitt, R and P Gottschalk (2002), "Trends in the Transitory Variance of Earnings in the United States" *Economic Journal*, Vol. 112, pp. 68-73.
- Neal, D and S Rosen (2000), "Theories of the Distribution of Earnings", in A Atkinson and F Bourguignon (eds.), *Handbook of Income Distribution*, Vol. 1, Elsevier Science B.V.
- Newey, W (1985) "Generalized Method of Moments Specification Testing" *Journal of Econometrics*, Vol. 29, pp. 229-256.
- Ramos, X (2003), "The Covariance Structure of Earnings in Great Britain, 1991-1999", *Economica*, Vol. 70, pp.353-374.
- Riksförsäkringsverket (1984), "Intjänad pensionspoäng inom ATP 1960-1982", Statistisk rapport 1984:7, Riksförsäkringsverket.
- Rosen, S (1997), "Public Employment, Taxes, and the Welfare State in Sweden", in R Freeman, R Topel and B Swedenborg (eds.), *The Welfare State in Transition*, The University of Chicago Press.
- Spånt, R (1979), "Den svenska inkomstfördelningens utveckling 1920-1976", background material to Långtidsutredningen 1978, Ds E 1979:4.
- Topel, R (1994), "Regional Labor Markets and the Determinants of Wage Inequality", *American Economic Review*, Vol. 82, pp.17-22.

## Appendix A: The estimation procedure

The parameters of the model in equations (5)-(8) are estimated by applying the minimum distance estimator of Chamberlain (1984). I first estimate  $\mu_{bt}$  in equation (1) with the sample mean log of earnings for the 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. This simple procedure adjusts for age, year, and cohort effects on average earnings.

For each of the 24 sample cohorts, born 1912/13 through 1958/59, I construct the sample auto-covariance matrix of  $y_{ibt}$ . For the 2 cohorts observed for the whole 27-53 age-span, born 1932/33 and 1934/35, these are 27x27 matrices; the matrices for other cohorts have smaller dimension. The distinct elements of each cohort's auto-covariance matrix is listed in a vector  $\hat{\mathbf{C}}_b$ , which contain 378=(27x28)/2 elements for the 2 cohorts observed for the whole 27-53 age-span and fewer elements for the other cohorts. The  $\hat{\mathbf{C}}_b$  vectors are stacked into an aggregate vector  $\hat{\mathbf{C}}$ , which contains a total of 4,188 distinct auto-covariance elements.

Let  $\mathbf{C}$  be the population analog to  $\hat{\mathbf{C}}$ . Express my model's moment restrictions as  $\mathbf{C} = f(\boldsymbol{\theta})$ , where  $\boldsymbol{\theta}$  is a vector that contains all the parameters of my model. The model in equations (5)-(8) then implies that the general variance element in  $\mathbf{C}$  is

$$(A1) \quad \text{Var}(y_{ibt}) = p_t^2 q_c^2 (\sigma_u^2 + \sum_a \sigma_{ra}^2) + \rho_t^2 \text{Var}(\varepsilon_{ib,t-1}) \\ + (2\rho_t \delta_t + \delta_t^2) \lambda_{t-1}^2 (\gamma_0 + \gamma_1(a-1) + \gamma_2(a-1)^2) + \lambda_t^2 (\gamma_0 + \gamma_1 a + \gamma_2 a^2),$$

and that the general auto-covariance element for  $y_{ibt}$  and  $y_{ib,t-1}$  is

$$(A2) \quad \text{Cov}(y_{ibt}, y_{ib,t-1}) = p_t p_{t-1} q_c^2 (\sigma_u^2 + \sum_{a-1} \sigma_{ra}^2) + \rho_t \text{Var}(\varepsilon_{ib,t-1}) \\ + \delta_t \lambda_{t-1}^2 (\gamma_0 + \gamma_1(a-1) + \gamma_2(a-1)^2),$$

and that the general auto-covariance element for  $y_{ibt}$  and  $y_{ib,t-s}$ , where  $s > 1$ , is



$$(A3) \quad \begin{aligned} Cov(y_{ibt}, y_{ib,t-s}) &= p_t p_{t-s} q_c^2 (\sigma_u^2 + \sum_{a-s} \sigma_{ra}^2) + \left( \prod_{t-s+1}^t \rho_t \right) Var(\varepsilon_{ib,t-s}) \\ &+ \left( \prod_{t-s+2}^t \rho_t \right) \delta_{t-s+1} \lambda_{t-s}^2 (\gamma_0 + \gamma_1(a-s) + \gamma_2(a-s)^2). \end{aligned}$$

In equations (A1)-(A3),  $\rho_t$  and  $\delta_t$  are restricted to differ between the period 1960-67 and 1968-90, but not within these periods.

With  $\mathbf{C} = f(\boldsymbol{\theta})$  specified, the vector  $\hat{\boldsymbol{\theta}}$  is chosen to minimize a distance function

$$(A4) \quad \mathbf{D} = (\hat{\mathbf{C}} - f(\hat{\boldsymbol{\theta}}))' \mathbf{W} (\hat{\mathbf{C}} - f(\hat{\boldsymbol{\theta}})),$$

where  $\mathbf{W}$  is a positive definite weighting matrix. The asymptotically optimal choice of  $\mathbf{W}$  is the inverse of any consistent estimator of the covariance matrix of  $\mathbf{C}$ . However, Altonji and Segall (1996) and Clark (1996) provide Monte Carlo evidence of potential serious finite sample bias in the estimate of  $\boldsymbol{\theta}$  with this approach. I therefore follow the practice 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), robust standard errors for  $\hat{\boldsymbol{\theta}}$  are obtained from the formula

$$(A5) \quad (\mathbf{G}'\mathbf{G})^{-1} \mathbf{G}'\mathbf{V}\mathbf{G}(\mathbf{G}'\mathbf{G})^{-1},$$

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 estimated covariance matrices of each  $\hat{\mathbf{C}}_b$  vectors.

## **Appendix B: The covariance structure with mean-approximated data**

Table B1 displays estimates of the variance of log earnings based on correct and top-coded earnings as well as mean-approximated earnings where the Pareto distribution and the lognormal distribution have been applied. The upper half of the table presents results for selected cohorts and the period 1968-71 where I have artificially top-coded earnings above a specific threshold. The threshold is the same for all cohorts each year, but it differs between

**Table B1: Variance of log earnings for various measures of earnings**

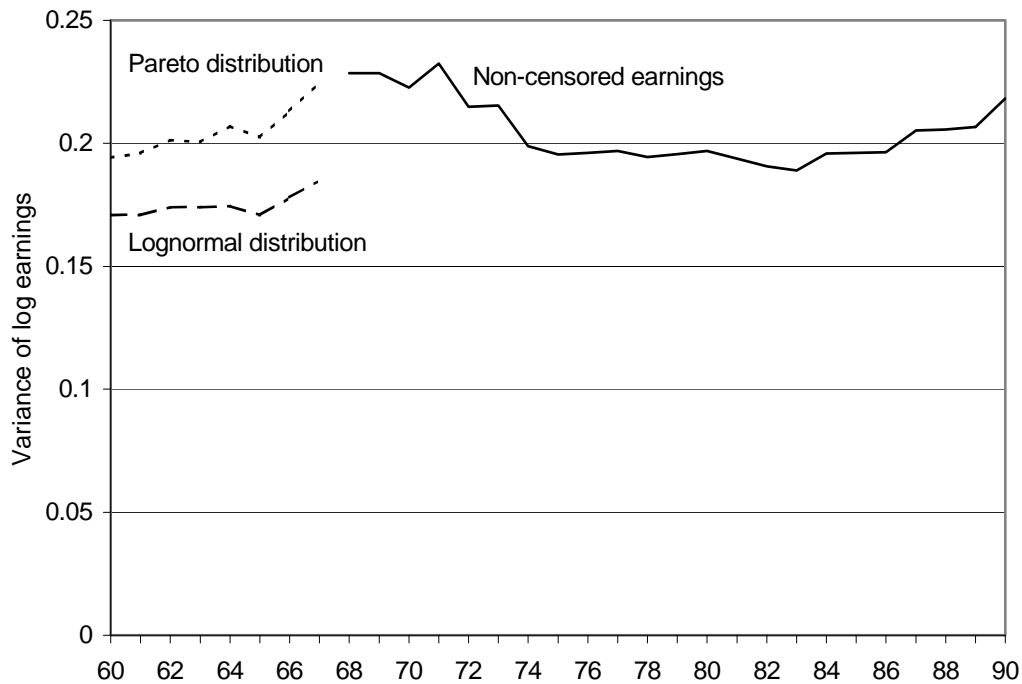
| <b>Artificially top-coded earnings<sup>a</sup></b> |             |                       |                         |   |  |
|--|-------------|-----------------------|-------------------------|---|--|
| Cohort born 1928/29                                |             |                       |                         |   |  |
| Year   | % top-coded | Var. correct earnings | Var. top-coded earnings | Var. mean-approx. earnings: Pareto distribution | Var. mean-approx. earnings: lognormal distribution |
| 1968   | 5.58        | 0.224                 | 0.191                   | 0.239   | 0.212  |
| 1969   | 10.08       | 0.241                 | 0.184                   | 0.253   | 0.214  |
| 1970   | 12.61       | 0.233                 | 0.171                   | 0.265   | 0.207  |
| 1971   | 13.38       | 0.246                 | 0.176                   | 0.246   | 0.218  |
| Cohort born 1934/35                                |             |                       |                         |   |  |
| 1968   | 3.06        | 0.188                 | 0.175                   | 0.192   | 0.187  |
| 1969   | 8.72        | 0.201                 | 0.169                   | 0.217   | 0.193  |
| 1970   | 10.78       | 0.202                 | 0.159                   | 0.215   | 0.189  |
| 1971   | 13.11       | 0.228                 | 0.176                   | 0.241   | 0.215  |
| Cohort born 1940/41                                |             |                       |                         |   |  |
| 1968   | 1.00        | 0.145                 | 0.140                   | 0.145   | 0.145  |
| 1969   | 3.00        | 0.157                 | 0.144                   | 0.155   | 0.150  |
| 1970   | 5.00        | 0.172                 | 0.150                   | 0.168   | 0.163  |
| 1971   | 7.00        | 0.182                 | 0.155                   | 0.178   | 0.172  |
| <b>Real top-coded earnings<sup>b</sup></b>         |             |                       |                         |   |  |
| Cohort born 1928/29                                |             |                       |                         |   |  |
| Year   | % top-coded | Var. correct earnings | Var. top-coded earnings | Var. mean-approx. earnings: Pareto distribution | Var. mean-approx. earnings: lognormal distribution |
| 1968   | 14.99       | 0.204                 | 0.135                   | 0.221   | 0.172  |
| 1969   | 16.99*      | 0.216                 | 0.139                   | 0.218   | 0.181  |
| 1970   | 19.48*      | 0.205                 | 0.121                   | 0.222   | 0.164  |
| 1971   | 19.48*      | 0.214                 | 0.133                   | 0.234   | 0.180  |
| Cohort born 1934/35                                |             |                       |                         |   |  |
| 1968   | 13.45       | 0.173                 | 0.130                   | 0.190   | 0.161  |
| 1969   | 16.60       | 0.186                 | 0.131                   | 0.203   | 0.169  |
| 1970   | 19.19*      | 0.185                 | 0.122                   | 0.199   | 0.164  |
| 1971   | 19.78*      | 0.194                 | 0.130                   | 0.228   | 0.176  |
| Cohort born 1940/41                                |             |                       |                         |   |  |
| 1968   | 4.04        | 0.137                 | 0.122                   | 0.133   | 0.130  |
| 1969   | 7.29        | 0.147                 | 0.125                   | 0.147   | 0.140  |
| 1970   | 11.14       | 0.159                 | 0.122                   | 0.154   | 0.144  |
| 1971   | 11.50       | 0.166                 | 0.128                   | 0.161   | 0.153  |

<sup>a</sup> All the variances are based on data from the IR.

<sup>b</sup> The last three columns is based on data from the PIR, and only individuals present in the PIR are included in the calculation of the variance for the correct earnings. The \* indicates that in any cohort the percentage censored is never this high between 1960 and 1967.

years in order to get the magnitude of top-coded observations to correspond to those typically present in the 1960-67 period. The lower part of Table B1 also presents some results based on original top-coded data from the PIR and the 1968-71 period. Here, the results for the correct data are based on earnings observations from the IR for the individuals present in the PIR. Note that the share of top-coded observations in the PIR is much higher during 1968-71 than during 1960-67.

The mean-approximation based on both distributions works well when the percentage top-coded is low; see the results for the cohort born 1940/41 based on artificial top-coding. Although both distributions to some extent underestimate the variances, the estimated movements over time are quite accurate, especially compared to those for the unadjusted top-coded data. The differences occur when the share top-coded is around 10 percent. The data based on the lognormal distribution still underestimates the true variance but the sign of the changes over time are generally correct, even though the magnitudes of the changes are too small. The data based on the Pareto distribution instead has a tendency to produce quite large overestimates of the variances when the percentage top-coded is high; for example, see the results based on artificially top-coded earnings for the cohort born 1934/35 and for the cohort born 1928/29 in 1970. This is a problem because the percentage top-coded in 1960 is much lower than that in 1967. The direction of the bias for the estimated variances will therefore probably differ between years and I do not know for which years I have overestimations. With the lognormal distribution, the direction of the bias is more certain as it during 1968-71 always underestimates the true variance; this is also true both for other cohorts and cutoff points than those presented in Table B1. One should, however, also be aware that the underestimations produced by the lognormal distribution sometimes are larger when the percentage top-coded is higher.



**Figure B1: The variance of log earnings for mean-approximated top-coded earnings 1960-67, based on the Pareto and lognormal distributions, respectively, and for non-censored earnings 1968-90.**

Note: Sample ages are used; see Table 1.

To further compare the mean approximations based on the Pareto and lognormal distributions, I have used these two distributions on each cohort in the 1960-67 sample and then pooled the cohorts. Figure B1 contains the variance of log earnings for these pooled samples together with the variance of log earnings during 1968-90 for the sample used in the main analysis. As expected, the variances produced by the Pareto distribution are larger than those produced by the lognormal distribution, but more important are the produced changes over time. The P80-P20 ratio of earnings and the variance of log earnings followed a very similar pattern during 1968-90. Based on the assumption that this also is true for the 1960-67 period, the variance should be quite constant. This is also the pattern produced by the lognormal distribution. The linear increase produced by the Pareto distribution is not in line with the evolution of the P80-P20 ratio. In fact, this linear increase seems very unlikely considered the equalizing wage policy at the time (see e.g. Hibbs, 1990; Edin and Topel, 1997).

**Table B2: Auto-correlation matrices for artificially top-coded earnings 1968-71, earnings 1972 are not top-coded.**

| Cohort born 1940/41 |                  |       |       |       |                    |       |       |       |  |       |       |       |   |       |       |       |
|---------------------|------------------|-------|-------|-------|--------------------|-------|-------|-------|--|-------|-------|-------|---|-------|-------|-------|
| Year                | Correct earnings |       |       |       | Top-coded earnings |       |       |       | Mean-approx. earnings: Pareto distribution |       |       |       | Mean-approx. earnings: lognormal distribution |       |       |       |
|                     | 1969             | 1970  | 1971  | 1972  | 1969               | 1970  | 1971  | 1972  | 1969                                       | 1970  | 1971  | 1972  | 1969  | 1970  | 1971  | 1972  |
| 1968                | 0.760            | 0.636 | 0.577 | 0.553 | 0.752              | 0.621 | 0.554 | 0.544 | 0.756                                      | 0.627 | 0.565 | 0.549 | 0.754   | 0.626 | 0.563 | 0.549 |
| 1969                |                  | 0.781 | 0.674 | 0.643 | 0.767              | 0.652 | 0.620 | 0.620 | 0.773                                      | 0.664 | 0.629 | 0.629 | 0.771   | 0.662 | 0.662 | 0.626 |
| 1970                |                  |       | 0.801 | 0.708 | 0.780              | 0.676 | 0.676 | 0.676 | 0.791                                      | 0.689 | 0.689 | 0.689 | 0.789   | 0.686 | 0.789 | 0.686 |
| 1971                |                  |       |       | 0.799 | 0.756              | 0.756 | 0.756 | 0.756 | 0.776                                      | 0.776 | 0.776 | 0.776 | 0.772   | 0.772 | 0.772 | 0.772 |
| Cohort born 1934/35 |                  |       |       |       |                    |       |       |       |  |       |       |       |   |       |       |       |
| Year                | 1969             | 1970  | 1971  | 1972  | 1969               | 1970  | 1971  | 1972  | 1969                                       | 1970  | 1971  | 1972  | 1969  | 1970  | 1971  | 1972  |
| 1968                | 0.857            | 0.799 | 0.756 | 0.713 | 0.845              | 0.777 | 0.721 | 0.709 | 0.850                                      | 0.789 | 0.750 | 0.716 | 0.849   | 0.787 | 0.744 | 0.715 |
| 1969                |                  | 0.849 | 0.794 | 0.740 | 0.831              | 0.764 | 0.717 | 0.717 | 0.847                                      | 0.787 | 0.736 | 0.736 | 0.842   | 0.783 | 0.783 | 0.732 |
| 1970                |                  |       | 0.856 | 0.801 | 0.835              | 0.770 | 0.770 | 0.770 | 0.851                                      | 0.793 | 0.793 | 0.793 | 0.848   | 0.788 | 0.848 | 0.788 |
| 1971                |                  |       |       | 0.870 | 0.824              | 0.824 | 0.824 | 0.824 | 0.856                                      | 0.856 | 0.856 | 0.856 | 0.850   | 0.850 | 0.850 | 0.850 |
| Cohort born 1928/29 |                  |       |       |       |                    |       |       |       |  |       |       |       |   |       |       |       |
| Year                | 1969             | 1970  | 1971  | 1972  | 1969               | 1970  | 1971  | 1972  | 1969                                       | 1970  | 1971  | 1972  | 1969  | 1970  | 1971  | 1972  |
| 1968                | 0.886            | 0.832 | 0.806 | 0.782 | 0.862              | 0.792 | 0.749 | 0.754 | 0.873                                      | 0.814 | 0.781 | 0.770 | 0.871   | 0.808 | 0.774 | 0.766 |
| 1969                |                  | 0.888 | 0.836 | 0.801 | 0.859              | 0.859 | 0.794 | 0.751 | 0.879                                      | 0.822 | 0.781 | 0.781 | 0.873   | 0.814 | 0.814 | 0.772 |
| 1970                |                  |       | 0.883 | 0.844 | 0.848              | 0.848 | 0.848 | 0.788 | 0.865                                      | 0.865 | 0.811 | 0.811 | 0.862   | 0.808 | 0.862 | 0.808 |
| 1971                |                  |       |       | 0.890 | 0.825              | 0.825 | 0.825 | 0.825 | 0.859                                      | 0.859 | 0.859 | 0.859 | 0.851   | 0.851 | 0.851 | 0.851 |

Note: All data is based on the IR.

**Table B3: Auto-correlation matrices for real top-coded earnings 1968-71, earnings 1972 are not top-coded.**

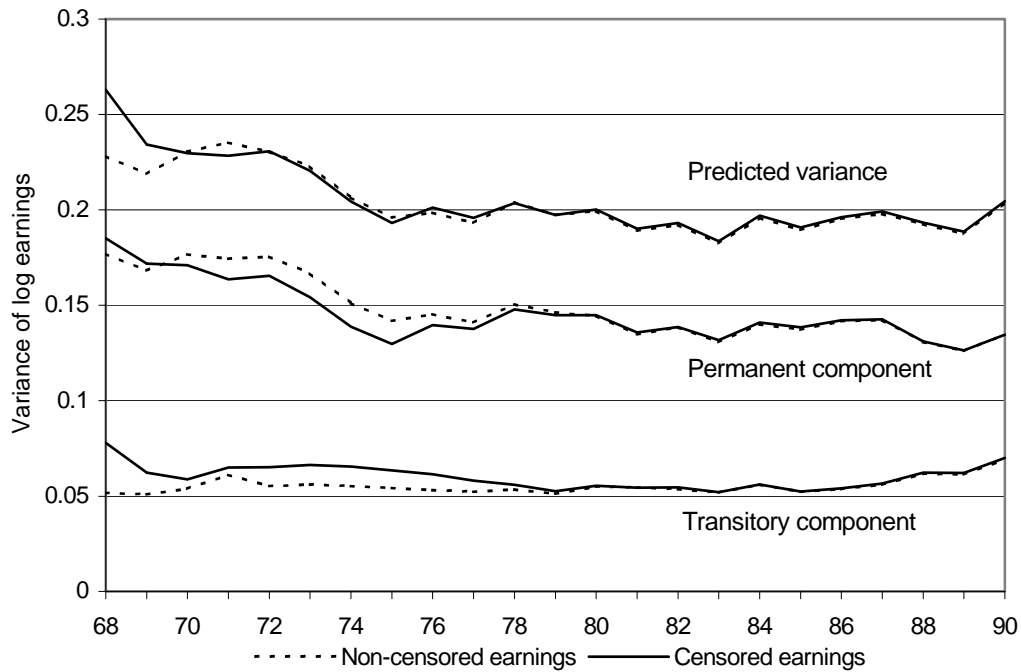
| Cohort born 1940/41 |                  |       |       |       |                    |       |        |       |  |       |        |       |   |       |       |       |
|---------------------|------------------|-------|-------|-------|--------------------|-------|--------|-------|--|-------|--------|-------|---|-------|-------|-------|
| Year                | Correct earnings |       |       |       | Top-coded earnings |       |        |       | Mean-approx. earnings: Pareto distribution |       |        |       | Mean-approx. earnings: lognormal distribution |       |       |       |
|                     | 1969             | 1970  | 1971  | 1972  | 1969               | 1970  | 1971   | 1972  | 1969                                       | 1970  | 1971   | 1972  | 1969  | 1970  | 1971  | 1972  |
| 1968                | 0.756            | 0.634 | 0.575 | 0.557 | 0.746              | 0.620 | 0.545  | 0.536 | 0.758                                      | 0.639 | 0.570  | 0.552 | 0.756   | 0.636 | 0.566 | 0.549 |
| 1969                |                  | 0.772 | 0.656 | 0.635 |                    | 0.752 | 0.618  | 0.584 |  | 0.772 | 0.647  | 0.607 |   | 0.768 | 0.642 | 0.603 |
| 1970                |                  |       | 0.791 | 0.696 |                    |       | 0.7575 | 0.640 |  |       | 0.7787 | 0.667 |   |       | 0.775 | 0.663 |
| 1971                |                  |       |       | 0.796 |                    |       |        | 0.729 |  |       |        | 0.761 |   |       |       | 0.757 |
| Cohort born 1934/35 |                  |       |       |       |                    |       |        |       |  |       |        |       |   |       |       |       |
| Year                | 1969             | 1970  | 1971  | 1972  | 1969               | 1970  | 1971   | 1972  | 1969                                       | 1970  | 1971   | 1972  | 1969  | 1970  | 1971  | 1972  |
| 1968                | 0.859            | 0.798 | 0.760 | 0.719 | 0.833              | 0.754 | 0.708  | 0.674 | 0.853                                      | 0.792 | 0.754  | 0.710 | 0.849   | 0.783 | 0.742 | 0.702 |
| 1969                |                  | 0.850 | 0.790 | 0.740 |                    | 0.823 | 0.744  | 0.686 |  | 0.855 | 0.790  | 0.725 |   | 0.847 | 0.777 | 0.718 |
| 1970                |                  |       | 0.859 | 0.808 |                    |       | 0.826  | 0.743 |  |       | 0.859  | 0.789 |   |       | 0.850 | 0.781 |
| 1971                |                  |       |       | 0.873 |                    |       |        | 0.785 |  |       |        | 0.831 |   |       |       | 0.825 |
| Cohort born 1928/29 |                  |       |       |       |                    |       |        |       |  |       |        |       |   |       |       |       |
| Year                | 1969             | 1970  | 1971  | 1972  | 1969               | 1970  | 1971   | 1972  | 1969                                       | 1970  | 1971   | 1972  | 1969  | 1970  | 1971  | 1972  |
| 1968                | 0.884            | 0.835 | 0.805 | 0.781 | 0.846              | 0.773 | 0.728  | 0.706 | 0.872                                      | 0.819 | 0.784  | 0.751 | 0.865   | 0.804 | 0.767 | 0.739 |
| 1969                |                  | 0.886 | 0.830 | 0.793 |                    | 0.840 | 0.780  | 0.711 |  | 0.870 | 0.820  | 0.763 |   | 0.862 | 0.809 | 0.752 |
| 1970                |                  |       | 0.880 | 0.837 |                    |       | 0.835  | 0.744 |  |       | 0.863  | 0.791 |   |       | 0.856 | 0.783 |
| 1971                |                  |       |       | 0.884 |                    |       |        | 0.782 |  |       |        | 0.834 |   |       |       | 0.825 |

Note: Earnings for 1972 are from the IR. The correct earnings are from the IR, and only individuals who also have earnings in the PIR are included. All other data 1968-71 are from the PIR.

Based on the same 1968-71 data as in Table B1 and non-censored earnings for 1972, Table B2 and B3 present estimated auto-correlation matrices for the 1968-72 period (non-censored earnings are included for 1972 to mimic the 1968 data). Studying Table B2 and B3 makes it clear that the use of the Pareto distribution works better than the lognormal distribution, but that the difference is small. The mean-approximated data based on both distributions generally produce auto-correlations slightly lower than the true values but with correct changes over time. It is also clear that the unadjusted top-coded earnings cause underestimates of the auto-correlations and thus overestimates earnings mobility; the same result is also reported by Hofer and Weber (2002) based on top-coded Austrian wage data.

### **Appendix C: Estimates for 1968-90 with top-coded 1968-75 earnings**

The results in Table 3 in the main text indicate that the permanent component decreased between 1960 and 1967 while the transitory component instead increased. Could this result be an artifact from the continuous increase in the share of top-coded earnings? Or is it in some way due to sample selection? To further investigate how the top-coded observations, sample selection, and the mean-approximation influence the results for the 1960-67 period, I have adjusted the accurate 1968-76 earnings data so as to replicate the top-coded 1960-68 data. First, only individuals included in the PIR between 1968 and 1976 are selected. This imposes sample selection as they cannot have pensions and must still be alive and living in Sweden in 1990 - although the selection is smaller than during 1960-68. Second, their recorded earnings from the IR during 1968-76 are artificially top-coded to match the top-coded 1960-68 earnings. That is, Table 1 shows that in 1960 the percentage top-coded is 4.65, the cut-off point for 1968 is then set so that the percentage top-coded this year is 4.65. The percentage top-coded earnings 1961 is 5.51, so the 1969 earnings data is transformed to having 5.51 percentage top-coded observations, and so forth until 1976 which corresponds to 1968 in the



**Figure C1: Decomposition of the variance of log earnings among men 40 years old, based on 1968-90 panels where one panel contains accurate earnings and the other contains mean-approximated top-coded earnings 1968-75.**

1960-90 analysis (the share top-coded in 1968 is 12.96 percent). The top-coded values for the 1968-75 period are then approximated by their estimated mean values based on the lognormal distribution for each cohort and year in the same manner as done for the 1960-67 data. To estimate the scale constant  $k_b$ , I use the top-coded observations with sample selection for 1976 and the accurate earnings without sample selection for the same year. This corresponds to the use of the 1968 data for the estimate of the scale constant for the 1960-67 earnings. Based on the mean-approximated top-coded 1968-75 earnings and accurate 1976-90 earnings, I estimate the model in equations (5)-(8) for the years 1968-90. Figure C1 contains the two predicted components for 40-year olds. As a benchmark, the results for the years 1968-90 from Figure 5, based on a 1968-90 panel with correct earnings data, are also included in Figure C1.

The mean-approximated data correctly capture the main pattern of the permanent component. For the transitory component, the mean-approximated data leads to constant



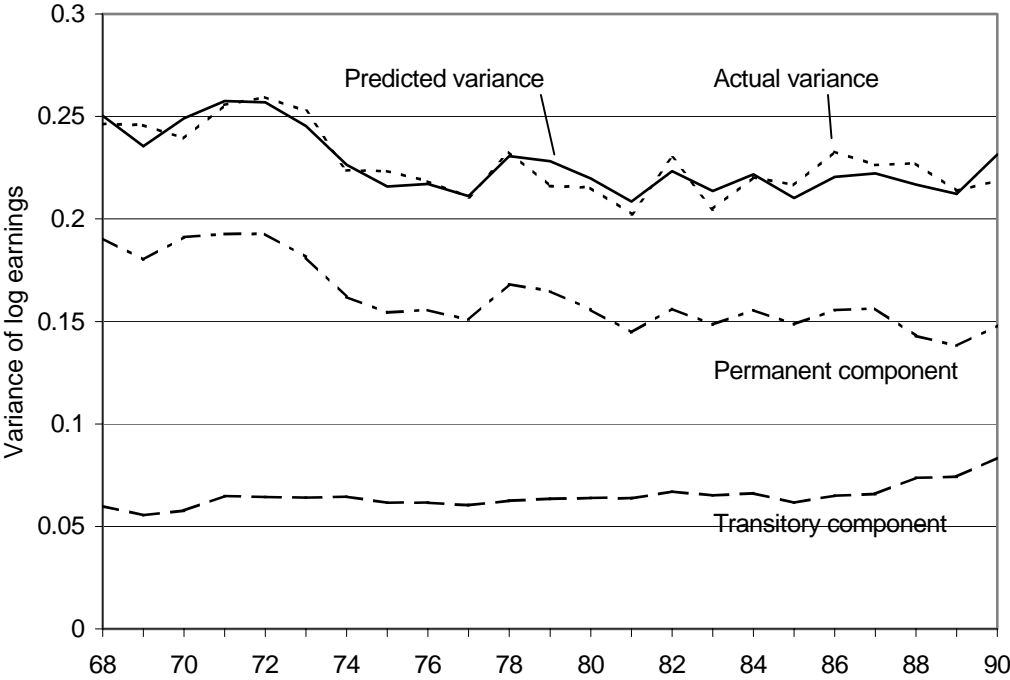
estimates during some of the years when the estimates based on accurate data instead decrease slightly; hence, there seems to be a slight effect from the year-to-year increase in the percentage of earnings that are top-coded. However, despite some wrong year-to-year movements for the transitory component, the estimates correctly capture the decrease during the 1970s as a whole.

For inference concerning the 1960-67 period, it should also be observed that the transitory component in Figure C1 is overestimated and that the overestimation is larger in the first two years. The permanent component is underestimated, except in the first two years when it is overestimated. These results indicate that the estimated scale parameters,  $k_b$ , are too large for the first sample years and that this causes the estimates of both components to be too large in these years. It is likely that this also applies to the results for the 1960-67 period. The estimated increase in the transitory component between 1960 and 1967 might thus actually be too small, as the transitory component is likely to be more overestimated in the beginning of the 1960s. In the same manner, the estimated decrease in the permanent component might be too large.

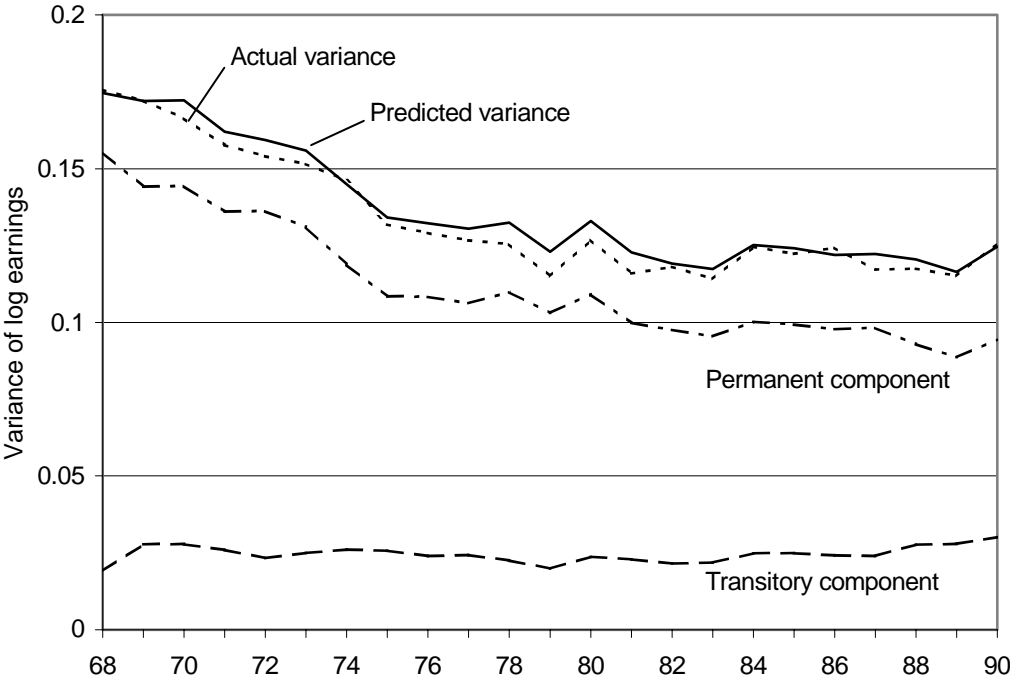
I have also performed these estimations based on data where the Pareto distribution has been applied to the top-coded earnings. These estimates were fairly accurate for the evolution of the permanent component, but produced wrong conclusions for the evolution of the transitory component during the 1970s. However, the estimates of changes in the relative size of the two components were correct (available on request).

I have also used top-coded data without sample selection, that is, without the requirement that an individual also has recorded earnings in the PIR. The results based on these data are very similar to those obtained with sample selection (available on request).

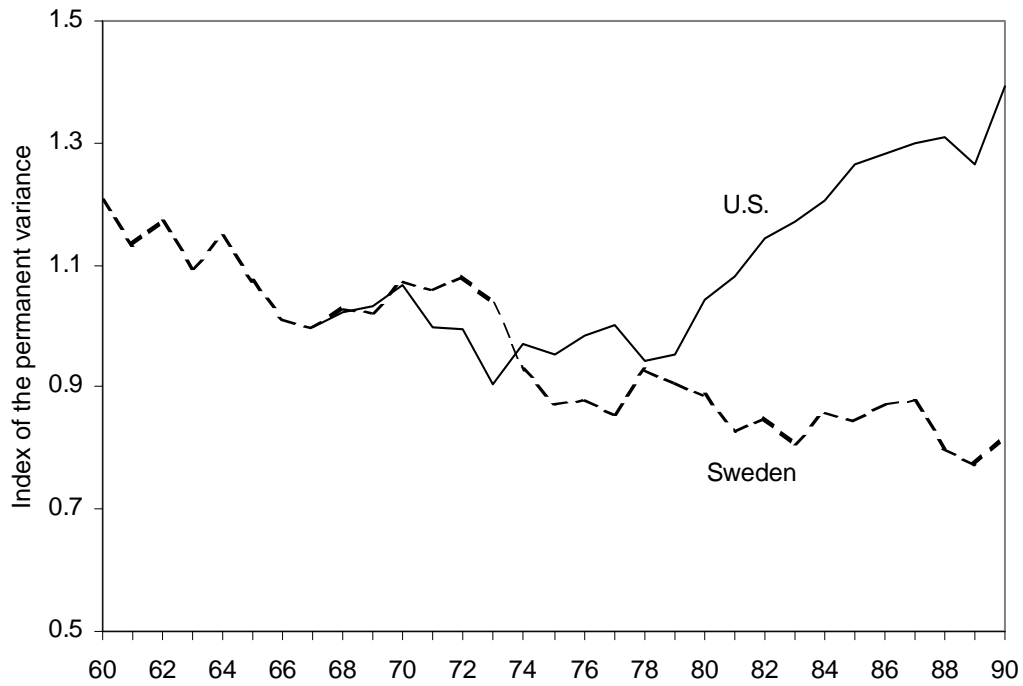
**Appendix D: Additional figures**



**Figure D1: Decomposition of the variance of log earnings among men 40 years old, based on a 1968-90 panel where the earnings threshold for including an individual in a given year is 27,595 real SEK in 2002.**



**Figure D2: Decomposition of the variance of log earnings among men 40 years old, based on a 1968-90 panel where the earnings threshold in a given year is the basic amount in 1960 adjusted for year-to-year changes in real wages.**



**Figure D3: The evolution of the permanent variance of log earnings in the U.S. between 1967 and 1990 and in Sweden between 1960 and 1990.**

Source: The numbers for the U.S. are based on Haider (2001), Table 5, column 2, p. 817. The numbers for Sweden are based on Table 3.

Note: The value for 1967 is normalized to unity in both series. The numbers for Sweden are for 40 year-olds and cohort heterogeneity is included in the predictions. Haider (2001) does not allow for cohort heterogeneity in the permanent component of earnings.

## WORKING PAPERS\*

Editor: Nils Gottfries

- 2003:10 Peter Fredriksson and Bertil Holmlund, Improving Incentives in Unemployment Insurance: A Review of Recent Research. 30 pp.
- 2003:11 Susanne Dahlberg and Ruth Aïda Nahum, Cohort Effects on Earnings Profiles: Evidence from Sweden. 47 pp.
- 2003:12 Matias Eklöf, Assessing Social Costs of Inefficient Procurement Design. 31 pp.
- 2003:13 Bertil Holmlund, The Rise and Fall of Swedish Unemployment. 52 pp.
- 2003:14 Karin Edmark, The Effects of Unemployment on Property Crime: Evidence from a Period of Unusually Large Swings in the Business Cycle. 27 pp.
- 2003:15 Patrik Hesselius, Does Sick Absence Increase the Risk of Unemployment? 32 pp.
- 2003:16 Mohammad Taslimi, Optimal Earnings-Related Unemployment Benefits. 30 pp.
- 2003:17 Peter Fredriksson and Bertil Holmlund, Optimal Unemployment Insurance Design: Time Limits, Monitoring, or Workfare? 25 pp.
- 2003:18 Yvonne Åberg, Peter Hedström and Ann-Sofie Kolm, Social Interactions and Unemployment. 36 pp.
- 2003:19 Tobias Lindhe, Jan Södersten and Ann Öberg, Economic Effects of Taxing Different Organizational Forms under a Dual Income Tax. 22 pp.
- 2003:20 Pär Österholm, The Taylor Rule – A Spurious Regression? 28 pp.
- 2003:21 Pär Österholm, Testing for Cointegration in Misspecified Systems – A Monte Carlo Study of Size Distortions. 32 pp.
- 2003:22 Ann-Sofie Kolm and Birthe Larsen, Does Tax Evasion Affect Unemployment and Educational Choice? 36 pp.
- 2003:23 Daniel Hallberg, A Description of Routes Out of the Labor Force for Workers in Sweden. 50 pp.
- 2003:24 N. Anders Klevmarken, On Household Wealth Trends in Sweden over the 1990s. 20 pp.
- 2003:25 Mats A. Bergman, When Should an Incumbent Be Obligated to Share its Infrastructure with an Entrant Under the General Competition Rules? 21 pp.

---

\* A list of papers in this series from earlier years will be sent on request by the department.

- 2003:26 Niclas Berggren and Henrik Jordahl, Does Free Trade Really Reduce Growth? Further Testing Using the Economic Freedom Index. 19 pp.
- 2003:27 Eleni Savvidou, The Relationship Between Skilled Labor and Technical Change. 44 pp.
- 2003:28 Per Pettersson-Lidbom and Matz Dahlberg, An Empirical Approach for Evaluating Soft Budget Constraints. 31 pp.
- 2003:29 Nils Gottfries, Booms and Busts in EMU. 34 pp.
- 2004:1 Iida Häkkinen, Working while enrolled in a university: Does it pay? 37 pp.
- 2004:2 Matz Dahlberg, Eva Mörk and Hanna Ågren, Do Politicians' Preferences Correspond to those of the Voters? An Investigation of Political Representation. 34 pp.
- 2004:3 Lars Lindvall, Does Public Spending on Youths Affect Crime Rates? 40 pp.
- 2004:4 Thomas Aronsson and Sören Blomquist, Redistribution and Provision of Public Goods in an Economic Federation. 23 pp.
- 2004:5 Matias Eklöf and Daniel Hallberg, Private Alternatives and Early Retirement Programs. 30 pp.
- 2004:6 Bertil Holmlund, Sickness Absence and Search Unemployment. 38 pp.
- 2004:7 Magnus Lundin, Nils Gottfries and Tomas Lindström, Price and Investment Dynamics: An Empirical Analysis of Plant Level Data. 41 pp.
- 2004:8 Maria Vredin Johansson, Allocation and Ex Ante Cost Efficiency of a Swedish Subsidy for Environmental Sustainability: The Local Investment Program. 26 pp.
- 2004:9 Sören Blomquist and Vidar Christiansen, Taxation and Heterogeneous Preferences. 29 pp.
- 2004:10 Magnus Gustavsson, Changes in Educational Wage Premiums in Sweden: 1992-2001. 36 pp.
- 2004:11 Magnus Gustavsson, Trends in the Transitory Variance of Earnings: Evidence from Sweden 1960-1990 and a Comparison with the United States. 63 pp.

See also working papers published by the Office of Labour Market Policy Evaluation  
<http://www.ifau.se/>