

Giesecke, Matthias; Bönke, Timm; Lüthen, Holger

**Conference Paper**

## The Dynamics of Earnings in Germany: Evidence from Social Security Records

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2011: Die Ordnung der Weltwirtschaft: Lektionen aus der Krise - Session: Empirical Labour Markets and Social Security, No. B11-V1

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Giesecke, Matthias; Bönke, Timm; Lüthen, Holger (2011) : The Dynamics of Earnings in Germany: Evidence from Social Security Records, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2011: Die Ordnung der Weltwirtschaft: Lektionen aus der Krise - Session: Empirical Labour Markets and Social Security, No. B11-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/48692>

**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.*

# The Dynamics of Earnings in Germany: Evidence from Social Security Records\*

Timm Bönke<sup>†</sup>  
Matthias Giesecke<sup>‡</sup>  
Holger Lüthen<sup>§</sup>

February 22, 2011

preliminary draft - please do not quote

## Abstract

We examine patterns of earnings volatility for male employees who are subject to statutory social security contributions in West Germany over the period 1986 - 2005. For this purpose, we analyse individual records covering highly reliable earnings biographies provided by the German Social Security Administration. We decompose earnings into permanent and transitory components and estimate parameters of the underlying variance-covariance structure of the earnings components model. This provides insights into the mechanics of earnings dynamics of the German labour market. We find evidence for increasing overall volatility which is predominantly driven by the permanent earnings component and therefore indicates increasing earnings inequality.

**Keywords:** Earnings Inequality, Earnings Dynamics, Variance Decomposition

**JEL-Classification:** C23, D31, J31

---

\*Lorenzo Cappellari is gratefully acknowledged for providing a basic programme on computing empirical autocovariances and minimum distance estimation.

<sup>†</sup>Institute for Public Finance, Freie Universität Berlin.

<sup>‡</sup>Ruhr Graduate School in Economics and Universität Duisburg-Essen:  
[Matthias.Giesecke@uni-due.de](mailto:Matthias.Giesecke@uni-due.de).

<sup>§</sup>Graduate Center, German Institute for Economic Research (DIW) Berlin.

# 1 Introduction

Economic uncertainty has been a well examined issue over the past few decades. The well-being of individuals in modern economies predominantly depends on their economic performance in terms of earnings. From an individual perspective, economic uncertainty can largely be ascribed to earnings related risk as to be induced by earnings volatility. Concerning Germany, empirical studies typically find a relatively stable earnings distribution over the 1980's (Wagner and Steiner, 1998). Starting in the mid-1990's, many studies find evidence for sharply rising earnings inequality (Gernandt and Pfeiffer, 2006). Much research has been conducted on issues dealing with cross-sectional earnings inequality. As a measure for one point in time, this proceeding cannot discover changes from one period to another. A cross-sectional analysis therefore does not allow for differentiations between permanent and transitory aspects of earnings dynamics. Thus, unraveling permanent and transitory earnings components is only possible in a longitudinal analysis and is particularly beneficial when evaluating governmental interventions to reduce earnings inequality.

The analysis of earnings volatility can be very useful to assess the evolution of earnings inequality. Increasing earnings related risk can reduce overall welfare, as noted by Blundell and Preston (1998). Under the assumption of risk aversion they find evidence that increasing short term risk can generate welfare costs on households since they might not be able to completely smooth income fluctuations. Moffitt and Gottschalk (2008) argue, that increasing earnings instability statistically contributes to rising cross-sectional variance and partially explains changes in cross-sectional inequality. If increasing earnings inequality is predominantly driven by persistent attributes such as education, then increasing cross-sectional inequality must be ascribed to inequality in longrun earnings. Contrarily, if a rise in cross-sectional inequality is rather driven by transitory components of earnings variation such as short term labour market developments, we are facing a year-to-year churning through the ranks of the annual earnings distribution and in such a case, long-run earnings inequality does not necessarily increase (see Baker and Solon, 2003). From a public finance perspective it is of particular interest to discuss the extent of welfare state interventions to reduce welfare costs or inequality caused by increasing earnings volatility. This study investigates the earnings covariance structure in Germany in search of earnings patterns with a main focus on aggregate change of earnings variation and life-cycle characteristics. Unraveling and understanding such patterns is necessary to contribute to the identification of instruments to reduce economic uncertainty in terms of earnings volatility. The evolution of earnings dynamics in Germany is assessed over a 20-year period, i.e. from 1986 to 2005. Earnings volatility is studied by exploring longitudinal earnings patterns of male employees, who are subject to social security contributions. For this purpose, an extraordinary dataset of the German Federal Pension Insurance provides process-produced earnings information of high reliability. Theoretical models allow to decompose earn-

ings into permanent and transitory components. Subsequently, the parameters of the underlying variance-covariance structure are estimated using Equally Weighted Minimum Distance (EWMD) estimation. To disentangle longrun and shortrun earnings components will appear to be useful, as the relative weight of transitory earnings fluctuations induced by short term phenomena can be analysed in comparison to life-cycle related and therefore permanent earnings heterogeneity. The results indicate increasing overall volatility which is predominantly driven by the permanent earnings component. We find evidence, that dispersion within the permanent earnings component seems to increase over calendar time - apart from the commonly known life cycle related earnings divergence. This indicates increasing earnings inequality.

Over the last three decades, a vast body of literature has developed on the covariance structure of longitudinal earnings data. Early contributions were made by Lillard and Weiss (1979), who examine sources of variation in earnings of American scientists in the 1960's. Hause (1980) provides a detailed discussion on Swedish earnings profiles with respect to the training on the job hypothesis. Early studies all have in common, that they apply maximum likelihood estimation techniques. Abowd and Card (1989) were the first ones to apply Chamberlains method of minimum distance estimation to the covariance structure of earnings.<sup>1</sup> Gottschalk and Moffitt (1994; 1995; 2008) have contributed influential studies covering data of the United States. More recent approaches by Cappellari (2004) on Italy and Baker and Solon (2003) on Canada both use administrative data instead of surveys. In recent years, some studies have covered the German case. Biewen (2005) examines changes in disposable household income inequality while Myck et al. (2008) focus on the explanation of changes in the relative importance of permanent and transitory earnings. In contrast to the present paper, both of these studies use German survey data.

The remainder of this study is structured as follows: *Section 2* explains the extraordinary nature of the earnings profiles from administrative data, describes sample restrictions and the birth cohort composition. Also, this section provides a discussion of Pareto-Imputation as a solution concept for missing earnings information due to top coding. *Section 3* introduces theoretical models for the variance-covariance structure of earnings and outlines the estimation procedure. *Section 4* provides results and illustrates them graphically. *Section 5* concludes.

## 2 Data

We use administrative data of the German statutory pension insurance, which is collected as individuals contribute to the pay-as-you-go pension system. The German pension system is based on the equivalence principle and complete earnings histories are documented

---

<sup>1</sup>See Chamberlain, 1984.

for each employee. Earnings histories are available only for those who are subject to social security contributions, but nevertheless the predominant part of the population is covered. Since data collection is part of an administrative process, the data quality is extraordinarily high. Besides earnings information, some socio-economic variables such as educational degrees are available. However, as the purpose for the data collection is to calculate pension insurance claims, only few variables on the household level exist and therefore our research on earnings volatility is restricted to the individual level. Advantages of social security data are the great reliability of earnings information since non-replies, incorrect answers or recall errors can be ruled out. In Germany, this kind of dataset is somewhat extraordinary and allows for an empirical analysis of yet unknown quality. Particularly in the research field of age-earnings profiles, process-produced data disclose a new dimension of possibilities (see Himmelreicher and Stegmann, 2008).

This study is based on a specific sample of insurance accounts, the so called *Insurant Account Sample (Versicherungskontenstichprobe)*. It originally served as basis for internal calculations of the German Federal Pension Insurance (*Deutsche Rentenversicherung Bund*) and for political consulting to support policy making. The Insurance Account Sample is a stratified random sample which includes a population of individuals who (i) live in Germany, (ii) have at least one entry in their individual insurance record and (iii) are aged between 15 and 67 in the reference year. The data collection is structured as a panel and the first reference year assembled is 1983.<sup>2</sup> The Insurant Account Sample is provided as scientific-use-file (SUF) for research purposes by the Data Research Center (*Forschungsdatenzentrum*) of the German Federal Pension Insurance for the reference years 2005, 2006 and 2007. In this present study, we use the SUF for the reference year 2006.<sup>3</sup> It is based upon a 25 % subsample of the original Insurant Account Sample. Due to the validation process of insurance records only individuals which are at least 30 in the respective reference year are included.<sup>4</sup> Altogether, the SUF contains 60,304 individual records in 2006. Available information encompass employment, unemployment, the type of pension insurance, periods of illness and others. Insurance records cover the biographies from the year insured persons reach age 14 until they turn 66. In sum, detailed information on 624 months are available. Special interest concerning this study relies on earnings information within the Insurant Account Sample. Earnings information consist of individual earnings histories, which are available over a long period of time. The longitudinal structure of person-based earnings information allows for a precise observation of changes in the earnings structure.

In spite of the extraordinary data quality, some limitations remain. First, administrative data are always subject to changes in legislation and hence, possible inconsistencies have to be corrected (Dustman et al., 2009). Second, selection takes places, as the Insurant

---

<sup>2</sup>For a detailed data documentation, see Deutsche Rentenversicherung Bund, 2008.

<sup>3</sup>This SUF ist officially referred to as SUFVSKT2006.

<sup>4</sup>Accordingly, reference year 2006 contains birth cohorts 1939 to 1976.

Account Sample only covers persons who are subject to the statutory pension insurance. Consequently, certain groups of the population such as self employed or civil servants are systematically excluded as long as they do not participate in the statutory pension insurance voluntarily. This is undoubtedly one source of bias and must be kept in mind while derivating results. Third, in the German statutory pension insurance exists an upper threshold of yearly maximum contributions, which is referred to as a contribution ceiling. For each year, a fixed income threshold delimits contributions to the pension scheme for those who possess very high earnings. Hence, earnings are top coded and the contribution ceiling inevitably leads to a maximum possible value of earnings points for each year. Since the analysis draws on earnings points to decode individual earnings, it is biased in a sense, that very high earnings above the yearly threshold are entirely assigned to the same value. Therefore, exact earnings cannot be decoded from earnings points for affected individuals and we face a tremendous loss of information at the top of the income distribution. In order to investigate earnings volatility, assumptions concerning earnings above the contribution ceiling have to be made. To obtain resilient information beyond the upper earnings threshold, Pareto imputation is implemented.

## 2.1 Pareto Imputation of Earnings for Top Coded Observations

The existence and variation of the contribution ceiling over time and the resulting differences in top coding of the annual earnings distribution are problematic. The existence of right censoring causes extensive restrictions in the analysis of earnings volatility and thus has to be accounted for. The contribution ceiling is subject to fundamental changes over time. For instance, in 1992 about 9.5 % out of West Germany's insured male population had earnings at or above the contribution ceiling. In 2003 only 6.95 % of the same underlying population were affected (see Bönke, 2009). To ensure time consistency and to enable the inclusion of the earnings distributions' upper tail in the analysis, top coded incomes must be imputed.

We use an imputation procedure for the data of the Insurant Account Sample proposed by Bönke (2009). The imputation of incomes for top coded observations assumes the upper tail of the income distribution to follow a Pareto-distribution. Several studies investigating income distributions in various countries indicate that the assumption of a Pareto-distributed upper tail of income distributions is a good approximation. For example, Piketty and Saez (2003) utilise a Pareto-based imputation method to study inequality in U.S. tax data. Kopczuk et al. (2009) base their analysis on data obtained from U.S. social security records which are top coded as well. They employ the parameters of the Pareto-distribution estimated by Piketty and Saez (2003) to impute earnings income for right censored observations. We use information of the earnings distribution as far as available to estimate parameters of the Pareto-distribution. Those parameters are subsequently used to impute earnings above the contribution ceiling, where observations at

or above the contribution ceiling must follow a rank criterion.<sup>5</sup> Comparing the obtained annual earnings distributions upper tail to (almost) uncensored survey based micro data reveals a good fit (see Bönke, 2009).

## 2.2 Sample Restrictions

To conduct an accurate analysis of variance, several sample restrictions are introduced. These come along with a stepwise sample reduction. Restrictions on the primary 60,304 observations within the Insurant Account Sample are necessary to obtain a suitable sample structure with respect to sample length, birth cohort structure and age structure.

The analysis of earnings volatility targets to cover the period between 1986 and 2005, i.e. 20 years. By setting this time frame, a rather long period is observed to be able to analyse earnings volatility over two decades from the mid-1980's up to the recent past in 2005.<sup>6</sup> Although data availability is given for individuals who gave birth in 1939 up to those who were born in 1976, not all of them can be used. Under the prerequisite of complete earnings information for a given time setting, each additional year to be analysed reduces cell size. Thus, one drawback is that sample length has to be traded off with sample size. To obtain full earnings profiles for the period under consideration, at first all person-year observations of individuals who were born before 1946 and after 1966 are excluded. Second, the whole sample is constrained to contain only persons aged between 20 and 59 years during the time frame of relevance. Selecting similar age limits is common practice in studies on earnings volatility (Gottschalk and Moffitt, 1994). By this means, it is attempted to capture as much earnings information on employees as possible. However, a disadvantage of selecting rather wide age limits is the inclusion of variation related to life cycle extremes. If such systematic variation is not accounted for appropriately, misleading interpretations might result. To avoid intrinsic variance, a number of additional sample restrictions have to be made. Hence, individuals of male gender, West-German origin and members of the generic pension insurance are taken into consideration exclusively. Earnings biographies of women are more volatile for reasons of maternity and while traditionally much less involved in the labour market compared to men, the number of employed women has grown rapidly since the 1980's (see Erlinghagen, 2004). Concerning the former separation of Germany, two diverse economic systems existed. This deeply affected individual earnings profiles far beyond the reunification in 1990 (see Hauser, 2003). Without further assumptions, West-German earnings profiles cannot simply be pooled with East-German ones. The much smaller group of East-Germans exhibited large scale changes during the economic integration process and are therefore excluded. Moreover, the sample is restricted to the generic pension insurance for reasons of formal differences

---

<sup>5</sup>See Bönke, 2009, for a detailed description of the imputation procedure.

<sup>6</sup>The observed period of 20 years is notably long compared to many similar studies. Lillard and Weiss (1979) cover 11 years (i.e. the decade of the 1960's) and Cappellari (2000) covers a period of 17 years. Myck et al. (2008) cover 13 years in a study using German survey-data of the GSOEP.

within insurance accounts of other insurance groups. These restrictions ensure a relatively homogeneous group on behalf of a proper analysis of variance.

### 2.3 Construction of Market Income

Monthly earnings are documented as earnings points by the German Federal Pension Insurance. Earnings points  $EGP_{it}$  at time  $t$  are obtained by relating individual earnings  $y_{it}$  of individual  $i$  at time  $t$  to average earnings  $\mu_t$  at time  $t$  in a corresponding period, which gives

$$EGP_{it} = \frac{y_{it}}{\mu_t} \quad (1)$$

Consequently, if a person - say in one year - earns the exact average amount among all social security contributors, this person receives exactly one earnings point.

To obtain annual individual earnings from the data, monthly earnings points are aggregated into annual earnings points by summing up for each individual in a first step. In a second step, annual earnings points are used to reverse the above operation in (1) by simply writing

$$y_{it} = EGP_{it} * \mu_t \quad (2)$$

and therefore obtain yearly individual gross earnings. In a third step, annual individual gross earnings are used to calculate market incomes. Those are needed to ensure independence from the design of the social security system. Individual market incomes are then deflated by German consumer prices. To avoid relational asymmetries of growing versus diminishing earnings, the deflated market incomes are logarithmised to finally obtain the basic earnings variable which is used in further analysis.

### 2.4 Birth Cohort Structure

Birth cohorts are constructed by joining different age groups to gain life-cycle related information. This proceeding helps to preserve cell size, as the number of observations would be rather small, if age groups were analysed separately by year of birth. The overall-sample is divided into three cohorts of individuals born between 1946 and 1966. To each cohort, its central age is assigned.<sup>7</sup> All birth cohorts can be observed over the full length of the relevant period from 1986 to 2005. Alternative approaches propose a birth cohort structure, which does not necessarily cover the entire period observed (see Baker and Solon, 2003). Such a framework is useful to maximise sample size because earnings biographies are not restricted to last over the entire sample length by definition but are rather allowed

---

<sup>7</sup>This gains importance within the technical implementation of Minimum Distance estimates in later sections.



to phase in and out of the sample during the observation period. However, concerning variation in the data, such an approach adds complexity when estimating parameters of more theoretical models and can cause problems for parameter identification. Thus, in the setting of this work, cohorts are observed only if they include social security contributors for the complete time frame. The resulting birth-cohort structure is summarised in Table 1.

Table 1: Birth Cohort Structure of the Sample.

Cohort	Year of Birth	Person-Year Obs.		Central Age	
		Total	Mean	Initial Year	Final Year
1	1946-1952	45059	2253	37	56
2	1953-1959	62041	3102	30	49
3	1960-1966	72192	3610	23	42

*Source:* Own calculations using SUFVSKT2006 data.

*Note:* To each seven-year birth cohort, its central age is assigned. For the oldest cohort born between 1946 and 1952, for instance, central age is coherent with the year of birth 1949.

The number of observations changes each year within birth cohorts due to unemployment spells or individuals who leave and enter the sample for other reasons.<sup>8</sup> As the numbers of observations within cohorts changes from year to year, Table 1 reports person-year observations. In terms of person-year observations we can see that younger cohorts are overrepresented. To get an idea of the average year-by-year cohort size, mean observations of the 20-year sample period are listed. As a description of the age structure, the central age of each birth cohort is reported for the initial year (1986) and the final year (2005) respectively.<sup>9</sup> The oldest cohort is observed with a central age of 37 in the initial year and 56 in the final year whereas the youngest group moves from an initial age of 23 to a final age of 42.

### 3 Conceptual Framework

Subsequently, theoretical models are introduced to decompose earnings into permanent and transitory components. The resulting earnings components model will allow to estimate parameters of the underlying covariance structure. This proceeding will provide a more formal explanation of the mechanics of earnings volatility.

<sup>8</sup>Dropping out or moving into the sample can happen for many reasons, for example change of job status from employee to self-employed, death etc.

<sup>9</sup>Please note, that assigned central ages of birth cohorts should not be mistaken with real individual age. The sample in fact covers ages between 20 and 59 year old individuals as described above.

### 3.1 The Canonical Earnings Components Model

A reasonable starting point to introduce such theoretical earnings models is the distinction between permanent and transitory earnings components in the sense of Friedman (1957).<sup>10</sup> More recent studies refer to this approach as the canonical variance-components model of earnings dynamics. In our setting  $w_{it}$  denotes adjusted log earnings of individual  $i$  in year  $t$  and  $k$  indicates lag length. By drawing on the longitudinal dimension of the Insurant Account Sample, this model then follows the expression

$$w_{it} = \mu_i + \nu_{it} \tag{3}$$

where

$$i = 1, \dots, N \quad t = 0, \dots, T - 1$$

and assumptions on first and second order moments are

$$\mu_i \sim (0, \sigma_\mu^2) \quad \nu_{it} \sim (0, \sigma_\nu^2) \quad E(\mu_i \nu_{it}) = 0$$

In this simple model, adjusted log earnings  $w_{it}$  are explained by (i) a permanent component  $\mu_i$  in the shape of an individual constant and (ii) a transitory component  $\nu_{it}$  in the shape of random deviations from it. The expected value of both earnings components is zero and no correlation occurs between permanent and transitory earnings. The variance-covariance structure as implied by the assumptions on second moments can be written as

$$\begin{aligned} E[w_{it}w_{i(t-k)}] &= E[(\mu_i + \nu_{it})(\mu_i + \nu_{i(t-k)})] = \\ &= E[\mu_i^2 + \mu_i\nu_{i(t-k)} + \mu_i\nu_{it} + \nu_{it}\nu_{i(t-k)}] = \\ &= \underbrace{E[\mu_i^2]}_{\sigma_\mu^2} + \underbrace{E[\mu_i\nu_{i(t-k)}]}_{=0} + \underbrace{E[\mu_i\nu_{it}]}_{=0} + \underbrace{E[\nu_{it}\nu_{i(t-k)}]}_{\sigma_\nu^2 \text{ (if } k=0)} \end{aligned}$$

Thus, under the assumption of serially uncorrelated earnings we finally obtain

$$E[w_{it}w_{i(t-k)}] = \begin{cases} \sigma_\mu^2 + \sigma_\nu^2 & \text{if } k = 0 \\ \sigma_\mu^2 & \text{otherwise} \end{cases}$$

However, this setting embodies a few shortcomings. Following the assumptions of the

---

<sup>10</sup>Friedman (1957), however, applied this income decomposition in a consumption framework.

simple model, one observes serially uncorrelated transitory earnings, i.e.  $E(\nu_{it}\nu_{i(t-k)}) = 0$  for  $k \neq 0$ . This becomes obvious when accounting for the fact, that transitory fluctuations only contribute to variances, whereas the longterm component contributes to both variances and covariances. The striking feature of this improbable assumption is that transitory fluctuations, say labour market shocks, completely disappear after each period. Another unrealistic implication of the above model is the constancy of the individual permanent component  $\mu_i$ . Such a framework does not allow for individual development over an earnings biography and does not come along with human capital approaches, suggesting some variation in individual ability over the life cycle. Such variation could be driven by skill acquisition via On-the-Job-Training or any other kind of experience (see Hause, 1980). To deal with the weaknesses of the canonical model of earnings dynamics in equation (3), the analytical framework must be extended by relevant aspects in order to achieve some improvement.

### 3.2 Individual Heterogeneity Growth within Permanent Earnings

As proposed by previous studies of the earnings covariance structure, the individual constant of the permanent earnings component  $\mu_i$  is supplemented by an additional parameter. The first ones who have applied such a framework were Hause (1977, 1980) and Lillard and Weiss (1979). Approaches of this kind are often referred to as individual heterogeneity growth (see Baker and Solon, 2003) or a random growth model (RG) (see Lillard and Weiss, 1979; Cappellari, 2004). The additional parameter within the permanent component accounts for earnings growth over the life-cycle. This parameter represents the effect of omitted variables which affect the growth of earnings, corresponding to *measurable* variables such as learning ability. As proxy for omitted variables it is feasible to use experience in this context. For this purpose, the age structure of the Insurant Account Sample is exploited, since resilient information on experience is not available.<sup>11</sup> The structure of the resulting RG model is of the form

$$w_{iat}^P = \mu_i + \gamma_i a_{it} \quad (4)$$

where the  $P$ -superscript denotes permanent earnings,  $a$  represents age where  $a = 0, \dots, A$  and  $\gamma_i$  is a growth parameter of individual earnings.<sup>12</sup> Assumptions on first and second order moments of the individual-specific coefficients are now given by

$$(\mu_i, \gamma_i) \sim [(0, 0); (\sigma_\mu^2, \sigma_\gamma^2, \sigma_{\mu\gamma})]$$

<sup>11</sup>Lillard and Weiss (1979) discuss potential identification problems which might appear when interpreting RG components as interactions of unobservables with either experience or time.

<sup>12</sup>Note, that age is measured as deviation from the minimum observed age within the sample. Within this work, this is equivalent to a minimum age (deviation) of 0 to a maximum age (deviation) of 19. For a similar proceeding, see Cappellari, 2000.

The resulting variance-covariance structure of the RG model follows the form

$$\begin{aligned}
E[w_{iat}^P w_{i(a-k)(t-k)}^P] &= E[(\mu_i + \gamma_i a_{it})(\mu_i + \gamma_i a_{i(t-k)})] = \\
&= E[\mu_i^2 + \mu_i \gamma_i a_{i(t-k)} + \mu_i \gamma_i a_{it} + \gamma_i^2 a_{it} a_{i(t-k)}] = \\
&= \underbrace{E[\mu_i^2]}_{\sigma_\mu^2} + \underbrace{E[\mu_i \gamma_i a_{i(t-k)}]}_{\sigma_{\mu\gamma} a_{t-k}} + \underbrace{E[\mu_i \gamma_i a_{it}]}_{\sigma_{\mu\gamma} a_t} + \underbrace{E[\gamma_i^2 a_{it} a_{i(t-k)}]}_{\sigma_\gamma^2 a_t a_{t-k}} = \\
&= \sigma_\mu^2 + \sigma_{\mu\gamma} a_{t-k} + \sigma_{\mu\gamma} a_t + \sigma_\gamma^2 a_t a_{t-k}
\end{aligned}$$

In this setting, on the one hand we look at the variance  $\sigma_\mu^2$ , which represents *time-invariant* heterogeneity. This time-invariant variance component represents constant heterogeneity such as an educational degree, for example. On the other hand we obtain the variance  $\sigma_\gamma^2$ , describing *time-varying* individual heterogeneity. This variance component represents changes over the life-cycle such as skill acquisition or the ability to learn. Both variances,  $\sigma_\mu^2$  and  $\sigma_\gamma^2$  are by definition positive. Moreover, a very important part of this enhanced model for the permanent earnings component is the covariance between the individual-specific coefficients  $\sigma_{\mu\gamma}$ , which can be positive or negative - depending on the overall development of individual earnings growth. A positive covariance  $\sigma_{\mu\gamma}$  implies that  $\sigma_\mu^2$  and  $\sigma_\gamma^2$  are positively related. Technically, we can say that higher time-invariant heterogeneity can be associated with higher time-variant heterogeneity in such a case. Thus, permanent earnings dispersion increases over time, if it is initially large and vice versa. In this case, persons who are better off at the beginning will experience faster earnings growth over the life cycle. Their growth path is steeper compared to individuals, whose primary endowment (e.g. education) is lower. Contrarily, a negative covariance  $\sigma_{\mu\gamma}$  implies that there must exist some negative relationship between  $\sigma_\mu^2$  and  $\sigma_\gamma^2$ . In such a case, permanent earnings dispersion decreases if it is initially large and vice versa. This finding would support a hypothesis, where initially low earnings dispersion is offset by faster earnings growth over the life-cycle and more mobility occurs within the entire earnings distribution. One could argue, that opportunity costs of higher education are payed off later in the shape of faster earnings growth. Hause (1980) uses such an argumentation of “compensatory mechanisms” in a training-on-the-job framework, where he illustrates a tradeoff between initial earnings and the slope of an individual earnings profile.<sup>13</sup>

By introducing a random walk specification, some authors take another step to either complement the theoretical modelling of the permanent earnings component (see Baker and Solon, 2003) or to present an alternative to the RG specification (see Cappellari, 2000). However, by comparing the goodness of fit between a random growth and a random walk specification, Cappellari (2000) finds that a random growth model provides a better description of life-cycle earnings dynamics. A random walk specification is not considered

<sup>13</sup>For a more detailed discussion of the OJT hypothesis in a similar context, see Hause (1980).

in this paper.

### 3.3 Serial Correlation within Transitory Earnings

Another somewhat unrealistic assumption which must be relaxed is the one of serially uncorrelated transitory earnings. In the setting of equation (3), no transitory earnings shock lasts for more than one period. Regarding the autocovariance function this would cause an abrupt drop from within-year variances to covariances. In such a setting, transitory variance would be reshuffled randomly every year and would not have any impact on subsequent periods. As common practice in the literature, serial correlation is introduced by some low order autoregressive process (AR). Some studies combine an AR process with a moving average process additionally to model short term deviations from covariances.<sup>14</sup> To introduce earnings shock persistence in an applicable way, we employ an autoregressive process of order one (AR(1)). This accounts for serial correlation in the transitory earnings component and can be written as

$$\nu_{it} = \rho\nu_{i(t-1)} + \epsilon_{it} \quad (5)$$

where assumptions regarding first and second order moments follow

$$\epsilon_{it} \sim (0, \sigma_\epsilon^2) \quad \nu_{i0} \sim (0, \sigma_{\nu 0}^2)$$

Resulting from the covariance structure of the AR(1) process in equation (5), three parameters of interest emerge. The initial variance  $\sigma_{\nu 0}^2$  measures volatility at the beginning of the sample, i.e.  $t = 0$  and  $k = 0$ , whereas the variance  $\sigma_\epsilon^2$  measures volatility in subsequent years ( $t > 0$ ). The autoregressive parameter  $\rho$  measures the persistence of earnings shocks. Within the autoregressive process, this parameter captures volatility from earlier periods by construction.

### 3.4 Relative Weight of Permanent and Transitory Earnings Variation

So far, no changes of the earnings variation over years are explicitly modelled. Such an exercise is performed in a next step, where the analysis is extended by year-specific factor loadings on both components. The resulting model takes the form

---

<sup>14</sup>For instance, Cappellari (2000) implements an ARMA(1,1) but experiences convergence problems in the full model, i.e. when additionally accounting for year- and cohort-specific factor loadings. For this reason he drops the moving average component and estimates an AR(1) process. Baker and Solon (2003) on the contrary, manage to implement an ARMA process.

$$w_{iat} = p_t w_{iat}^P + \lambda_t \nu_{it} \quad (6)$$

where  $p_t$  is a time-shifter on the permanent component and  $\lambda_t$  is a time-shifter on the transitory component. These year-specific factor loadings account for aggregate changes in the earnings covariance structure. Particularly, time-shifters determine the relative weight of permanent and transitory earnings variation over time. In this setting, factor loadings of the two respective components could be interpreted as “prices” (Cappellari, 2000). From a cross-sectional perspective it is important to note, that factor loadings on permanent and transitory earnings components influence total variance in any given period  $t$ . Changes in total variance strongly depend on which one of the two components increases or decreases. Changes in  $p_t$  can be interpreted as changes in unobserved individual-specific permanent components such as skill acquisition, which have been modelled in the RG framework. On the other hand, changes in  $\lambda_t$  can be addressed to variation in short-term labour market volatility or institutional changes such as regulatory reforms.<sup>15</sup> The results to be discussed in the next section will highlight the importance of a distinction between aggregate changes within the two earnings components over time.

### 3.5 Birth Cohort Heterogeneity

To distinguish age effects and life-cycle effects, separate covariance matrices have been estimated for each birth cohort. Thus, empirical autocovariance functions can be generated for each birth cohort to observe earnings volatility for different stages of the life-cycle at the same point in time. The earnings components model presented so far must now be extended to account for such birth cohort heterogeneity. This is achieved by introducing cohort-specific dummy variables to allow for a separate identification of respective elements within the autocovariance function. Hence, (6) can be rewritten in the form

$$w_{iat} = \kappa_c p_t w_{iat}^P + \tau_c \lambda_t \nu_{it} \quad (7)$$

where  $\kappa_c$  is the birth cohort dummy for the permanent component and  $\tau_c$  for the transitory component respectively. The subscript  $c$  denotes central age of cohorts and is used to identify cohort-related elements of the empirical variance-covariance matrix within the empirical implementation. Equation (7) summarises the complete model as it embodies a decomposition of earnings into permanent and transitory components and accounts for year-specific factor loadings as well as birth cohort heterogeneity.

---

<sup>15</sup>Institutional changes, however, can affect earnings permanently, if such reforms exist for a long period.

### 3.6 Minimum Distance Estimation

To estimate relevant parameters of the specified models, Minimum Distance Estimation (MD) is employed as an application of the Generalised Method of Moments (GMM). Chamberlain (1984) provides a general discussion on GMM whereas applications of MD to earnings dynamics models can be found in a variety of studies.<sup>16</sup> In this work, MD is used to fit the parameters of the theoretical models to the data and allows to conduct an estimation without imposing assumptions regarding the earnings distribution.

The estimation procedure follows two preliminary steps. In the first step, the constructed market incomes (see section 2) are adjusted for age and calendar time in an ordinary least squares regression on year dummies and age. In the second step, adjusted earnings from the first stage regression are used to estimate variance-covariance matrices for each birth cohort separately. This is accomplished by estimating individual sample averages of cross products of earnings. For this purpose, earnings vectors for each year are constructed, which include adjusted earnings on the individual level.<sup>17</sup> Then, products of all vectors of adjusted earnings are taken, i.e. for all possible pairs of years. This yields  $T(T + 1)/2$  new vectors containing the products, where  $T$  is the number of years observed. Individual products are now summed up for each combination of years and sums are weighted by the existing number of individual earnings observations for according pairs of years. This yields  $T(T + 1)/2$  elements of the empirical variance-covariance matrix. As this study covers 20 years from 1986 to 2005, we obtain  $20(20 + 1)/2 = 210$  elements for the pooled sample, which are stacked in vector  $m$  afterwards. Formally, let  $M_c$  be an estimate of the  $T \times T$  autocovariance matrix of birth cohort  $c$ . Then a vector

$$m_c = Vech(M_c) \tag{8}$$

can be defined, which collects all  $(T(T + 1)/2)$  distinct elements of the empirical autocovariance matrix  $M_c$ . To find evidence on birth cohort related volatility, empirical variance-covariance matrices are not derived from the pooled sample but for each birth cohort separately. This proceeding finally yields  $C(T(T + 1)/2)$ , where  $C$  is the number of cohorts. Thus, we obtain  $3(20(20 + 1)/2) = 630$  elements for all three cohorts. Those 630 distinct elements are then stacked cohort by cohort, to generate a vector containing 630 variances and covariances. By constituting the dependent variable, these autocovariances within vector  $m$  play an important role in the subsequent estimation procedure.

---

<sup>16</sup>See Abowd and Card (1989) for a more formal explanation of the application of MD to earnings dynamics models. More recent studies with related applications are Biewen (2005), Cappellari (2004), Baker and Solon (2003), or Myck et al. (2008).

<sup>17</sup>Note, that a balanced sample is used in a sense, that individuals are not allowed to have temporarily missing values on earnings information.

Let  $f(\theta)$  define a non-linear function of the theoretical autocovariance structure, which embodies the parameters of interest. In the particular theoretical framework which has been specified within this section,  $\theta$  contains three parameters of the RG model for the permanent earnings component and another three parameters for the transitory component in the shape of an AR(1) process.<sup>18</sup> Additionally,  $\theta$  includes all calendar time shifters  $p_t$  and  $\lambda_t$  as well as birth cohort shifters  $\kappa_c$  and  $\tau_c$ . To obtain a consistent estimator of  $\theta$ , the squared distance between the empirical covariance structure  $m$  and the theoretical covariance structure  $f(\theta)$  must be minimised in the form

$$\hat{\theta} = \arg \min_{\theta} [m - f(\theta)]' W [m - f(\theta)] \quad (9)$$

Much discussion has arisen concerning the choice of the weighting matrix  $W$  in the minimisation step. Chamberlain (1984) shows that under fairly general conditions,  $W = V^{-1}$  is the optimal choice of a weighting matrix, where  $V^{-1}$  is the inverse of the earnings fourth moments matrix. The fourth moments matrix itself is a measure of the kurtosis and describes the distribution of probability-compound around the expected value of a random variable.<sup>19</sup> This, in fact, allows for an estimation without any distributional assumption on the underlying variable. The weighting procedure in the sense of Chamberlain (1984) provides at least asymptotic efficiency and is therefore called Optimal Minimum Distance. However, Altonji and Segal (1996) propose a more simple proceeding. They find evidence, that correlation between the sampling errors of second and fourth moments in Optimal Minimum Distance could lead to biased parameter estimates. To avoid these problems, they propose the identity matrix as weighting instrument. The resulting estimation procedure is commonly referred to as Equally Weighted Minimum Distance (EWMD), and is similar to a nonlinear least squares procedure (see, for example, Cappellari, 2004; Baker and Solon, 2003; Myck et al. 2008). Hence, the squared distance between the empirical covariance structure and its theoretical counterpart is minimised by estimating least squares on the non-linear function  $f(\theta)$ .<sup>20</sup> We use EWMD to estimate relevant parameters of the earnings covariance structure.

In terms of integrity, the theoretical covariance structure of the specified models as embodied in  $f(\theta)$  must be merged into one term. This is achieved by working out second moments as implied by the model in equation (7). The overall variation is decomposed into an RG model for the permanent variation and an AR(1) process for transitory variation

---

<sup>18</sup>The variance of constant individual heterogeneity  $\sigma_{\mu}^2$ , the variance of earnings growth  $\sigma_{\gamma}^2$  and the covariance  $\sigma_{\mu\gamma}$  respectively for the permanent component as well as the initial variance  $\sigma_{v0}^2$ , the year-to-year variance  $\sigma_{\epsilon}^2$  and the persistence parameter of the AR(1) process  $\rho$  for the transitory component.

<sup>19</sup>Third and fourth moments are usually used to measure the deviation of a given distribution from the normal distribution.

<sup>20</sup>Common statistical packages provide least squares estimation procedures for non-linear equations.



with birth cohort and calendar time factor loadings on both components, which gives

$$\begin{aligned}
f(\theta) &= E[w_{iat}w_{i(a-k)(t-k)}] = \\
&= \underbrace{\kappa_c \left[ p_t p_{t-k} E[w_{iat}^P w_{i(a-k)(t-k)}^P] \right]}_{\text{Permanent Component}} + \underbrace{\tau_c \left[ \lambda_t \lambda_{t-k} E[\nu_{it} \nu_{i(t-k)}] \right]}_{\text{Transitory Component}}
\end{aligned}$$

This gives an additively composed covariance structure, which is simply the sum of the covariance structure of the permanent and the transitory earnings component. For the covariance structure of the permanent component, this gives

$$E[w_{iat}^P w_{i(a-k)(t-k)}^P] = \sigma_\mu^2 + a_t a_{(t-k)} \sigma_\gamma^2 + (a_t + a_{(t-k)}) \sigma_{\mu\gamma}$$

and for the transitory component

$$E[\nu_{it} \nu_{i(t-k)}] = \begin{cases} \sigma_{\nu_0}^2 & \text{if } t = 0, k = 0 \\ \rho^2 \sigma_{\nu_0}^2 + \sigma_\epsilon^2 & \text{if } t = 1, k = 0 \\ \rho^2 E[\nu_{i(t-1)} \nu_{i(t-1)}] & \text{if } t > 1, k = 0 \\ \rho^k E[\nu_{i(t-k)} \nu_{i(t-k)}] & \text{if } k \leq t, 1 \leq k \leq T - 1 \end{cases}$$

respectively.

## 4 Results

This section summarises the results obtained from EWMD estimation on the Insurant Account Sample. For this purpose, 630 distinct elements of empirical variance-covariance matrices for three birth cohorts were used. Due to separate calculation of empirical variances and covariances for each birth cohort, the structure of the dataset allows to distinguish cohort specific and time-specific effects. The full model contains three sets of parameter estimates as specified in the previous section for (i) the theoretical models of the covariance structure, (ii) calendar time effects and (iii) birth cohort effects.

The first model of consideration includes estimates for the theoretical models of the covariance structure and birth cohort heterogeneity. The results are presented in Table 2.

The upper left part of the table shows results for the RG specification of the permanent earnings component. The results refer to the oldest cohort which has been normalised for identification. The estimated parameter for constant individual heterogeneity  $\sigma_\mu^2$  indicates some dispersion with respect to the initial earnings situation. The coefficient for heterogeneity growth  $\sigma_\gamma^2$  indicates some variation in earnings growth. From the parameter value of  $\hat{\sigma}_\gamma^2 = 0.0003$  we can infer, that an individual one standard deviation above the mean of

Table 2: EWMD Estimates including Birth Cohort Effects.

	Permanent Component			Transitory Component		
	coef.	s.e.		coef.	s.e.	
<i>Theoretical Models</i>						
$\sigma_\mu^2$	.1074	(.0072)	***	$\sigma_{\nu 0}^2$	.0163	(.0143)
$\sigma_\gamma^2$	.0003	(.0000)	***	$\sigma_\epsilon^2$	.0780	(.0045) ***
$\sigma_{\mu\gamma}$	.0015	(.0005)	**	$\rho$	.6551	(.0167) ***
<i>Birth Cohorts</i>						
$\kappa_1$ (1946-1952)	1.0000			$\tau_1$	1.0000	
$\kappa_2$ (1953-1959)	1.0445	(.0242)	***	$\tau_2$	1.1474	(.1348) ***
$\kappa_3$ (1960-1966)	0.7597	(.0225)	***	$\tau_3$	1.5328	(.1641) ***
<i>Model Fit</i>						
SSR Unweighted				0.0517		
SSR Weighted				7861.88		
N Observations				630		

*Source:* Own calculations using SUFVSKT2006 data. *Note:* \*\*\* p < 0.01; \*\* p < 0.05; \* p < 0.10. Equally Minimum Distance Estimates of the earnings components model on  $20(20+1)/2 = 630$  empirical moments. For identification, the parameters of the oldest cohort (1946-1952), i.e.  $\kappa_1$  and  $\tau_1$ , have been set to unity.

the distribution of  $\gamma$  shows earnings growth rates which are roughly 1.7% higher compared to the population mean within the oldest cohort.<sup>21</sup> As expected, some dispersion within individual earnings growth rates exists. The estimated covariance between intercept ( $\sigma_\mu^2$ ) and slope ( $\sigma_\gamma^2$ ),  $\sigma_{\mu\gamma}$ , is positive. This indicates a positive relationship between initial earnings dispersion and the life cycle evolution of earnings profiles. In this context, initially large earnings dispersion leads to even larger time varying heterogeneity and vice versa. Cappellari (2004) obtains similar findings but many studies have found negative signs (see for example Lillard and Weiss, 1979; Hause, 1980; Baker and Solon, 1998). In this case, compensatory mechanisms in the sense of Hause (1980) could be responsible for lower initial earnings dispersion to be traded off against steeper earnings growth paths and vice versa. This ambivalence seems to stem from comparisons of studies for different countries and different datasets.

The upper right part in Table 2 contains estimation results for the transitory earnings component. The variances  $\sigma_{\nu 0}^2$  and  $\sigma_\epsilon^2$  are needed for consistent estimation of the AR(1)-process.<sup>22</sup> The coefficient of the autoregressive process  $\rho$  is a measure of persistence of earnings shocks and shows a moderate value of  $\hat{\rho} = 0.66$  compared to other studies.<sup>23</sup> This

<sup>21</sup>Since  $\hat{\sigma}_\gamma^2 = 0.0003$  is the estimated (mean) variance for earnings growth rates, the standard deviation is given by  $\sqrt{\hat{\sigma}_\gamma^2} = 0.017$ . Reference group is the oldest cohort normalised to one.

<sup>22</sup>In usual time series approaches, an infinite past is assumed for AR(1) processes. MaCurdy (1982) however states, that such an assumption is not tenable when applying a time series process to disaggregated (individual) panel data. Assuming an infinite past is then unrealistic such that  $\sigma_{\nu 0}^2$  must be estimated for the initial period.

<sup>23</sup>In a study using the German Socio-Economic Panel (GSOEP), Myck et al. (2008) obtain a parameter value of  $\hat{\rho} = 0.62$  in their earnings specification and  $\hat{\rho} = 0.57$  for wages. In a study of earnings dynamics

implies that after five years, roughly 12% of an earnings shock to the transitory component are on average still present. This finding might indicate some moderate earnings stability in Germany but, notwithstanding, parameter values of such an autoregressive process can largely alternate as different model specifications are taken into account (see Baker and Solon, 2003). The persistence parameter might well depend on sample selection which happens to be the case in the Insurant Account Sample being restricted to employees only. Ignoring influential groups of the working population such as self-employed individuals could be a driving force for lower shock persistence, i.e. greater stability. For Germany, this aspect gains even more importance with respect to relatively well-organised (though declining) unions, which contribute to the overall stabilisation of earnings via wage bargaining (see Fachinger, 1991).<sup>24</sup>

The remainder of Table 2 consists of parameter estimates regarding the birth cohort structure. Parameters for the variation across birth cohorts have been estimated for both the permanent and the transitory earnings component, i.e.  $\kappa_c$  and  $\tau_c$ . Two central results shed light on the nature of volatility in earnings profiles over the life cycle. First, variation across birth cohorts with respect to the permanent earnings component is clearly lower as younger birth cohorts are taken into consideration. Second, the exact opposite picture surrenders with respect to the transitory earnings component and variation across birth cohorts is much larger as younger birth cohorts are observed. With regard to the degree of earnings dispersion, these are two diametrically opposed trends. For the permanent earnings component this indicates diverging earnings profiles. Thus, for a given point in time, different cohorts find themselves at different stages of their earnings curriculum and cohort effects suggest that older cohorts will exhibit higher dispersion in the context of permanent earnings aspects. The exact opposite picture emerges for cohort effects on the shortterm component, where younger cohorts clearly exhibit higher transitory fluctuations. Hence we observe that (i) older cohorts exhibit higher earnings differentials due to permanent earnings volatility while (ii) younger cohorts exhibit much more transitory volatility.

In presence of heteroskedasticity and autocorrelation, standard errors and conventional p-values cannot be fully trusted for statistical significance. To account for tenable inference, the bottom part of Table 2 reports the sum of squared residuals as a measure for the goodness of fit regarding a correct model specification. The sum of squared residuals is weighted by the inverse of the estimated variance of residuals. This measure is  $\chi^2$ -distributed with  $C(T(T+1)/2) - P$  degrees of freedom, where P is the number of esti-

---

on Canada, Baker and Solon (2003) report  $\hat{\rho} = 0.54$  in their base model but obtain much higher values of  $\hat{\rho} = 0.68$  and  $\hat{\rho} = 0.72$  when estimating additionally restricted models. Covering Italy, Cappellari (2000) even obtains larger values but those drop markedly when specifying a model on white collar workers only.

<sup>24</sup>German unions, however, have experienced large-scale attrition in the past two decades and therefore lost much power (see Müller-Jentsch and Itterman, 2000).

mated parameters.<sup>25</sup> The null-hypothesis of correct model specification is rejected under conventional significance levels. Hence, the weighted sum of squared residuals is used as a measure of fit, where the fitting performance increases as the measure is decreases.<sup>26</sup>

In Table 3, estimation results on aggregate changes of earnings volatility over time are documented.

Table 3: EWMD Estimates including Calendar Time Effects.

	Permanent Component			Transitory Component		
	coef.	s.e.		coef.	s.e.	
<i>Theoretical Models</i>						
$\sigma_{\mu}^2$	.1110	(.0043)	***	$\sigma_{\nu 0}^2$	.0727	(.0256) ***
$\sigma_{\gamma}^2$	.0002	(.0001)	***	$\sigma_{\epsilon}^2$	.1586	(.0219) ***
$\sigma_{\mu\gamma}$	.0014	(.0003)	***	$\rho$	.4221	(.0225) ***
<i>Years</i>						
$p_{1987}$	1.1506	(.0347)	***	$\lambda_{1987}$	.6449	(.1333) ***
$p_{1988}$	1.2526	(.0370)	***	$\lambda_{1988}$	.6971	(.1454) ***
$p_{1989}$	1.4115	(.0399)	***	$\lambda_{1989}$	.6058	(.1225) ***
$p_{1990}$	1.5107	(.0441)	***	$\lambda_{1990}$	.5334	(.1167) ***
$p_{1991}$	1.6445	(.0491)	***	$\lambda_{1991}$	.5070	(.1031) ***
$p_{1992}$	2.0229	(.0543)	***	$\lambda_{1992}$	.5673	(.0730) ***
$p_{1993}$	2.2479	(.0628)	***	$\lambda_{1993}$	.5151	(.0711) ***
$p_{1994}$	2.4748	(.0709)	***	$\lambda_{1994}$	.5339	(.0721) ***
$p_{1995}$	2.4581	(.0783)	***	$\lambda_{1995}$	.3492	(.0664) **
$p_{1996}$	2.7368	(.0885)	***	$\lambda_{1996}$	.3031	(.0658) **
$p_{1997}$	3.0345	(.1018)	***	$\lambda_{1997}$	.2972	(.0665) **
$p_{1998}$	3.3701	(.1168)	***	$\lambda_{1998}$	.2842	(.0714) **
$p_{1999}$	3.7062	(.1283)	***	$\lambda_{1999}$	.2697	(.0680) **
$p_{2000}$	4.3219	(.1443)	***	$\lambda_{2000}$	.3142	(.0747) **
$p_{2001}$	4.7282	(.1589)	***	$\lambda_{2001}$	.3650	(.0704) **
$p_{2002}$	5.2249	(.1738)	***	$\lambda_{2002}$	.4086	(.0726) **
$p_{2003}$	5.5476	(.2039)	***	$\lambda_{2003}$	.5417	(.0796) **
$p_{2004}$	5.5649	(.2217)	***	$\lambda_{2004}$	.7353	(.0919) **
$p_{2005}$	6.3732	(.3015)	***	$\lambda_{2005}$	.3851	(.1374) **
<i>Model Fit</i>						
SSR Unweighted				0.1193		
SSR Weighted				19565.12		
N Observations				630		

Source: Own calculations using SUFVSKT2006 data.

Note: \*\*\*  $p < 0.01$ ; \*\*  $p < 0.05$ ; \*  $p < 0.10$ . Equally Weighted Minimum Distance Estimates of the earnings components model on 630 empirical moments. For identification, the parameters of the first year of observation (1986), i.e.  $p_t$  and  $\lambda_t$ , have been set to unity.

The parameter estimates for the theoretical models vary only marginally and reflect the same principal results as in the model specification with birth cohort effects. The autoregressive parameter value of  $\hat{\rho} = 0.42$  is lower compared to the first model specification and confirms a low shock persistence among German employees. The lower part of Table 3 lists parameter estimates for the relative weight of permanent and transitory earnings components. Calendar time effects with respect to the permanent earnings component  $p_t$  constantly rise over the sample period and increase sharply towards the year 2005. By

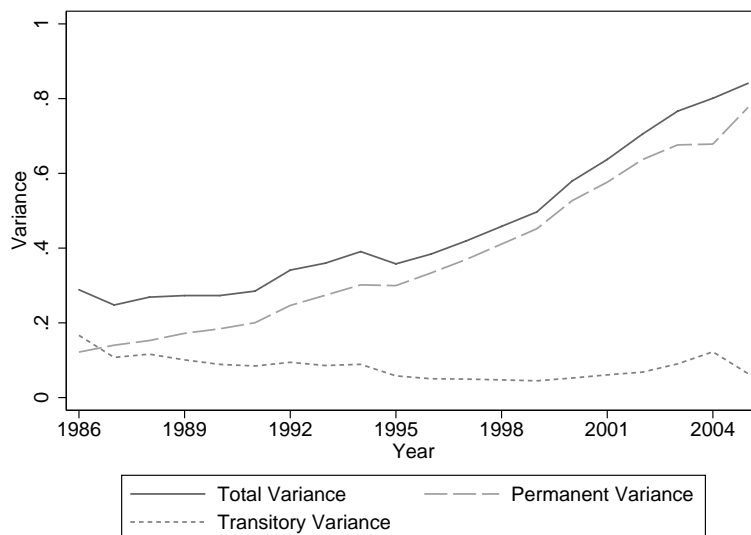
<sup>25</sup>In this model, this gives  $3(20(20+1)/2)-10=620$  degrees of freedom.

<sup>26</sup>The fitting performance is useful most notably in a comparison of different model specifications. However, this paper does not provide alternative model specifications which is a matter of future research.

contrast, calendar time effects for the transitory earnings component  $\lambda_t$  play a more important role at the beginning but decline slightly thereafter. Generally, transitory volatility is relatively stable over the observed 20 year period.

So far we have only discussed the absolute values of parameter estimates for both earnings components. However, it is of particular interest to get insights into the composition of permanent and transitory variance with respect to total variance. Predicted variances are decomposed in Figure 1, following the estimated covariance structure which results from equation (6) in *section 3*. This allows for a graphical illustration of changes in the volatility of respective earnings components.

Figure 1: Variance Decomposition for the Pooled Sample



*Source:* Own calculations using SUFVSKT2006 data.

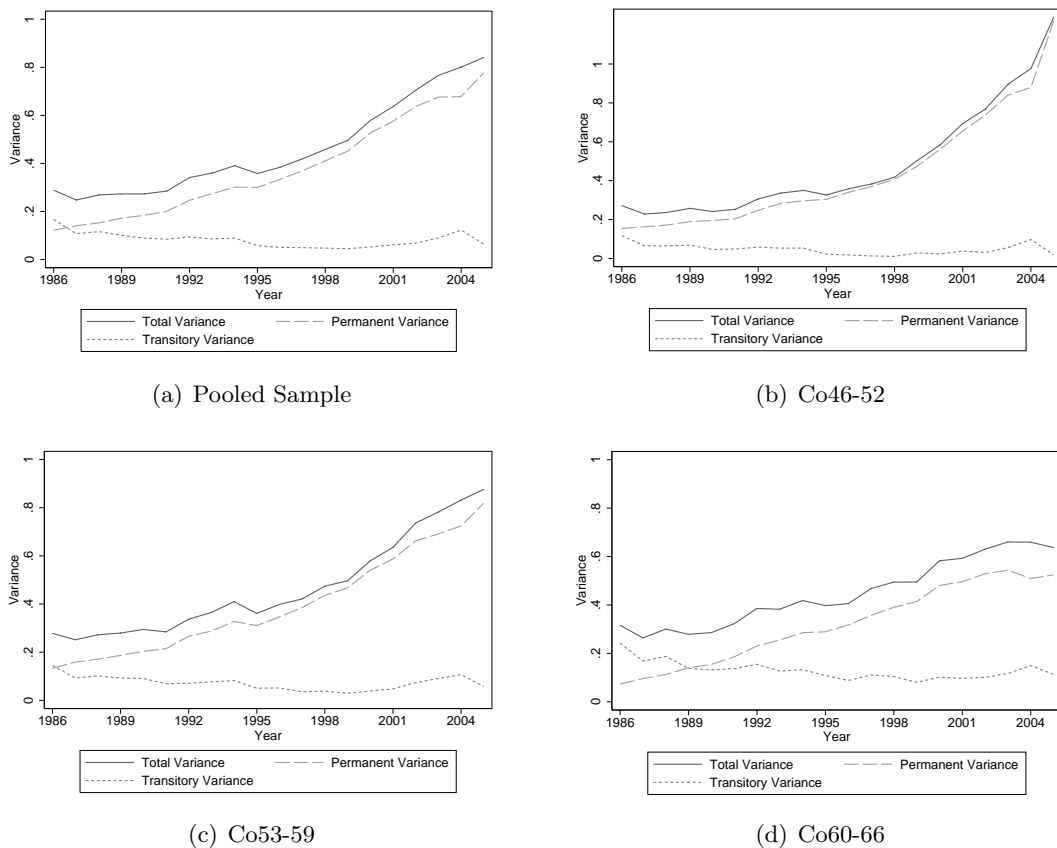
*Note:* Decomposition of predicted total variance into permanent and transitory components.

In terms of the relative weight of earnings components, transitory variance has a much larger share on total variance in the mid 1980's. However, towards the 1990's a reverse picture appears and permanent earnings variance gains an increasingly larger share on total variance. The decline of the transitory variance takes place smoothly while the variation in the permanent component increases remarkably, such that permanent and transitory variation diverge continuously. Figure 1 shows how the relative importance of permanent earnings variance has constantly increased between 1986 and 2005. This does not support findings for a decreasing fraction of permanent earnings variance in Germany as argued by Myck et al. (2008). Our findings indicate that the total variance is strongly driven by permanent earnings dispersion while transitory fluctuations remain relatively stable. As transitory volatility vanishes after few years and permanent volatility can im-

pact the earnings distribution and household-poverty (Cappellari, 2000), this finding is of utmost importance. Increasing permanent earnings dispersion is an unambiguous sign for increasing earnings inequality. We find this process of increasing earnings inequality in Germany for the two decades of observation, which is coherent to Bach et al. (2007) who find moderately increasing inequality between 1992 and 2001.

To gain further insight into life cycle related aspects, Figure 2 displays a variance decomposition separately for each cohort. All cohorts roughly resemble the patterns of the pooled sample. Transitory fluctuations seem to be slightly more pronounced for younger cohorts. At the beginning of the sample period in the mid 1980's, transitory variance is larger than permanent variance. Subsequently, permanent earnings differentials play a much larger role. A general characteristic of earnings volatility in the observed time frame is, without exception, the constant increase of variance within the permanent earnings component over time; regardless of which cohort is taken into consideration.

Figure 2: Variance Decomposition by Birth Cohort

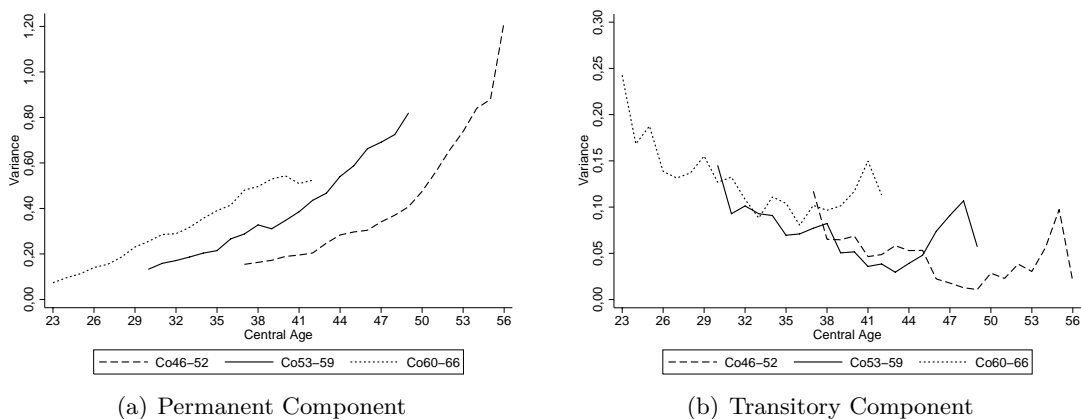


*Source:* Own calculations using SUFVSKT2006 data.

*Note:* Decomposition of predicted total variance into permanent and transitory components. Decomposition for pooled sample in (a) and for separate cohorts in (b)-(d).

So far we have found evidence for strongly increasing overall variance of earnings which is predominantly driven by dispersion within the permanent earnings component. Moreover, we found evidence for diverging earnings profiles over the life cycle. However, the latter finding is not very surprising as occupational differences and other kinds of heterogeneity naturally lead to some degree of divergence. Figure 3 shows that divergence within earnings profiles is not exclusively due to life cycle related aspects. This is established by plotting the evolution of the two variance components separately by central age for each of the three cohorts. As central age is depicted for each seven year birth cohort, the lines are staggered.

Figure 3: Variance Decomposition by Components and Birth Cohorts



*Source:* Own calculations using SUFVSKT2006 data.

*Note:* Decomposition of predicted total variance into permanent and transitory components in (a) and (b) respectively. Cohorts are depicted with their central age.

Subfigure (a) in Figure 3 shows the dispersion within the permanent earnings component for the three cohorts at a given central age. We observe that for a given central age, permanent earnings dispersion is much higher for younger cohorts. This finding suggests a strong link to calendar time, indicating a trend of increasing aggregate inequality in the sense of increasing permanent earnings dispersion. Clearly, increasing permanent earnings dispersion is a question not only of natural divergence over earnings profiles as it generally occurs over the life cycle. Subfigure (b) depicts transitory volatility by central age for the three cohorts. Transitory fluctuations are decreasing, although there is a peak at the end, which corresponds to the year 2004. This is due to a policy change with respect to minor employment. Contrasting the permanent case, we do not find supporting evidence for any trend in transitory volatility. Given an increasing total variance, both subfigures in Figure 3 support our finding of sharply increasing permanent earnings volatility over time which is not exclusively due to natural divergence as individuals proceed in their earnings biography.

## 5 Conclusion

This study examines earnings volatility of West-German men who are subject to social security contributions between 1986 and 2005. For this purpose, earnings histories from administrative data of the German Federal Pension Insurance are used. Pension insurance records provide highly reliable process-produced data. In a first step, theoretical models are used to decompose earnings into permanent and transitory components. The theoretical model allows to estimate parameters of the underlying covariance structure of earnings. Thereby, central parameters are supplemented by additional ones concerning aggregate changes over time and birth cohort heterogeneity. Those additional parameters allow to reveal some clear patterns of earnings volatility. To disentangle permanent and transitory earnings and their respective covariance structure is indispensable in search of explanations for increased total earnings volatility and related questions of earnings inequality.

Equally Weighted Minimum Distance estimates of the theoretical models have established a profound perspective on earnings dynamics by distinguishing permanent and transitory earnings components. Estimates of Random Growth parameters indicate constant individual heterogeneity and individual heterogeneity growth concerning the permanent earnings component. A positive covariance between the two suggests a positive relation between initial earnings dispersion and earnings growth dispersion and thus indicates increasing dispersion if the initial dispersion is large and vice versa. The persistence parameter within the transitory earnings component reflects a moderate earnings shock persistence. However, it must be kept in mind that model specification and the type of dataset used for estimation are driving forces of this precise result.

Effects on birth cohort heterogeneity provide useful insights into life cycle related aspects of earnings volatility. Two diametrically opposed effects of earnings variation have occurred. First, we find higher dispersion for older cohorts regarding the permanent earnings component. Second, we find that younger cohorts exhibit much more transitory earnings volatility. By setting this scene, it is not unambiguous to state whether earnings volatility increases with age or not. The answer to this question depends on the type of earnings component to which volatility refers to. The results suggest that short-term uncertainty is much larger initially and successively diminishes over the life cycle which seems to reflect difficulties when entering the labour market. Contrarily, relevant individual characteristics of permanent nature such as educational degrees are responsible for increasing earnings inequality over the life-cycle as divergence in earnings profiles takes place. But such divergence is not surprising whatsoever. We found clear evidence for a remarkable increase of the overall earnings variance during the observed period from 1986 to 2005. However, the variance decomposition shows that on the one hand, transitory earnings fluctuations have



remained on a stable level from the mid-1980's until 2005. On the other hand, dispersion of the permanent earnings component has increased dramatically and is the driving force of an overwhelming part of increasing earnings volatility. In this manner, permanent earnings dispersion has gained absolute as well as relative importance over several years. It is of utmost importance that increasing permanent earnings dispersion is not exclusively due to life cycle related divergence (see Figure 3 (a)). We rather find a clear trend of growing permanent earnings dispersion as calendar time proceeds. This finding is of such particular importance, because the longterm earnings component reflects returns on persistent worker attributes such as education and other primary endowment. Thus, an aggregate demand shift towards more skilled labour seems to take place, especially in a world with technological progress. Increasing permanent earnings differentials are of structural nature and therefore they do not simply vanish after few years. In such a situation, welfare state interventions must be considered. However, the instruments of policy makers should not be constrained to redistributive arrangements such as taxes and transfers to deal with the consequences. Much more than before, questions of educational policy as well as demographic change must be discussed to identify fundamental solutions.

## References

- [1] Abowd, John M. and David Card (1989): *On the Covariance Structure of Earnings and Hours Changes*, *Econometrica* 57(2), 411-445.
- [2] Altonji, Joseph G. and Lewis M. Segal (1996): *Small-Sample Bias in GMM Estimation of Covariance Structures*, *Journal of Business & Economic Statistics* 14(3), 353-366.
- [3] Bach, Stefan, Corneo, Giacomo and Viktor Steiner (2007): *From Bottom to Top: The Entire Distribution of Market Income in Germany*, Discussion Paper 683, DIW Berlin.
- [4] Baker, Michael and Gary Solon (2003): *Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records*, *Journal of Labor Economics* 21(2), 289-321.
- [5] Biewen, Martin (2005): *The Covariance Structure of East and West German Incomes and its Implications for the Persistence of Poverty and Inequality*, *German Economic Review* 6, 445-469.
- [6] Blundell, Richard and Ian Preston (1998): *Consumption Inequality and Income Uncertainty*, *Quarterly Journal of Economics* 113(2), 603-640.
- [7] Bönke, Timm (2009): *Gekappte Einkommen in prozessgenerierten Daten der Deutschen Rentenversicherung: Ein pareto-basierter Imputationsansatz*, in *Deutsche Rentenversicherung Bund* (eds.): *Bericht vom sechsten Workshop des Forschungsdatenzentrums der Rentenversicherung*, DRV-Schriften 55.
- [8] Cappellari, Lorenzo (2000): *The Dynamics and Inequality of Italian Male Earnings: Permanent Changes or Transitory Fluctuations?*, ISEER Working Paper, University of Essex.
- [9] Cappellari, Lorenzo (2004): *The Dynamics and Inequality of Italian Men's Earnings: Long-Term Changes or Transitory Fluctuations?*, *Journal of Human Resources* 39(2), 475-499.
- [10] Chamberlain, Gary (1984): *Panel Data*, *Handbook of Econometrics*, Zvi Griliches and Michael D. Intriligator (eds.), North Holland.
- [11] Dustman, Christian and Ludsteck, Johannes and Uta Schönberg (2009): *Revisiting the German Wage Structure*, *Quarterly Journal of Economics*, 843-881.
- [12] Erlinghagen, Marcel (2004): *Die Restrukturierung des Arbeitsmarktes: Arbeitsmarktmobilität und Beschäftigungsstabilität im Zeitverlauf*, Verlag für Sozialwissenschaften, Wiesbaden.
- [13] Fachinger, Uwe (1991): *Lohnmobilität in der Bundesrepublik Deutschland. Eine Untersuchung auf Basis der prozess-produzierten Längsschnittdaten der gesetzlichen Rentenversicherung*, Duncker und Humblot, Berlin.
- [14] Friedman, Milton (1957): *A Theory of the Consumption Function*, Princeton University Press, Princeton.

- [15] Gernandt, Johannes and Friedhelm Pfeiffer (2006): *Rising Wage Inequality in Germany*, ZEW Discussion Paper 06-019, ZEW Mannheim.
- [16] Gottschalk, Peter and Robert Moffitt (1994): *The Growth of Earnings Instability in the U.S. Labour Market*, Brookings Papers on Economic Activity 2, 217-254.
- [17] Hause, John C. (1977): *The Covariance Structure of Earnings and the On-the-Job Training Hypothesis*, Annals of Economic and Social Measurement 6(4), 6-38.
- [18] Hause, John C. (1980): *The Fine Structure of Earnings and the On-the-Job Training Hypothesis*, Econometrica 48(4), 1013-1029.
- [19] Himmelreicher, Ralf K. and Michael Stegmann (2008): *New Possibilities for Socio-Economic Research through Longitudinal Data from the Research Data Centre of the German Federal Pension Insurance (FDZ-RV)*, Schmollers Jahrbuch 128, 647-660.
- [20] Kopczuk, Wojciech and Saez, Emmanuel and Jae Song (2010): *Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937*, The Quarterly Journal of Economics 125, 91-128.
- [21] Lillard, Lee A. and Yoram Weiss (1979): *Components of Variation in Panel Earnings Data: American Scientist 1960-70*, Econometrica 46(5), 437-454.
- [22] Moffitt, Robert and Peter Gottschalk (1995): *Trends in the Covariance Structure of Earnings in the U.S.: 1969-1987*, Working Papers in Economics 355, The John Hopkins University.
- [23] Moffitt, Robert and Peter Gottschalk (2008): *Trends in the Transitory Variance of Male Earnings in the U.S., 1970-2004*, Boston College Working Papers in Economics 697, The Boston College.
- [24] Müller-Jentsch, Walther and Peter Ittermann (2000): *Industrielle Beziehungen. Daten, Zeitreihen, Trends 1950-1999*, Campus Verlag, Frankfurt am Main.
- [25] Myck, Michal, Ochmann, Richard and Salmai Qari (2008): *Dynamics of Earnings and Hourly Wages in Germany*, Discussion Paper 3751, IZA Bonn.
- [26] Müller, Kai U. and Viktor Steiner (2008): *Would a Legal Minimum Wage Reduce Poverty? A Microsimulation Study for Germany*, Discussion Paper 791, DIW Berlin.
- [27] Piketty, Thomas and Emmanuel Saez (2003): *Income Inequality in the United States, 1913-1998*, Quarterly Journal of Economics 118, 1-39.