



## SONDERFORSCHUNGSBEREICH 504

Rationalitätskonzepte,  
Entscheidungsverhalten und  
ökonomische Modellierung

No. 07-05

### **Lifetime Earnings and Life Expectancy**

Hans-Martin von Gaudecker\*  
and Rembrandt Scholz\*\*

February 2006

Financial support from the Deutsche Forschungsgemeinschaft, SFB 504, at the University of Mannheim, is gratefully acknowledged.

\*Mannheim Research Institute for the Economics of Aging, email: [HMGaudecker@web.de](mailto:HMGaudecker@web.de)

\*\*Max Planck Institute for Demographic Research, email: [scholz@rostockerzentrum.de](mailto:scholz@rostockerzentrum.de)



Universität Mannheim  
L 13,15  
68131 Mannheim

# Lifetime Earnings and Life Expectancy <sup>\*</sup>

Hans-Martin von Gaudecker <sup>†</sup>

Mannheim Research Institute for the Economics of Aging (MEA)

Rembrandt D. Scholz

Max Planck Institute for Demographic Research, Rostock

February 15, 2006

## Abstract

We estimate remaining life expectancy at age 65 using a very large sample of male German pensioners. Our analysis is entirely nonparametric. Furthermore, the data enable us to compare life expectancy in eastern and western Germany conditional on a measure of socio-economic status. Our findings show a lower bound of almost fifty percent (six years) on the difference in remaining life expectancy between the lowest and the highest socio-economic group considered. Within groups, we find similar values for East and West. Our analysis contributes to the literature in several aspects. First, Germany is clearly underrepresented in differential mortality studies. Second, we are able to use a novel measure of lifetime earnings as a proxy for socio-economic status that remains valid for retired people. Third, the comparison of eastern and western Germany may provide some interesting insights for transformation countries.

---

<sup>\*</sup>A substantial part of this research was undertaken while the first author was visiting the Forschungsdatenzentrum of the Deutsche Rentenversicherung (Research Data Centre of the German Pension Insurance), Berlin, in September and October of 2005. Thanks for invaluable help with the data and many important comments are due to Ralf Himmelreicher, Jonas Radl, Frank Röder, and Michael Stegmann. We are very grateful for comments from Vladimir Shkolnikov and seminar participants at the Deutsche Rentenversicherung Bund that greatly helped to improve upon the paper.

<sup>†</sup>*Address for correspondence:* Mannheim Research Institute for the Economics of Aging; L 13, 17; 68131 Mannheim; Germany. Email: gaudecker@mea.uni-mannheim.de

# 1 Introduction

The international literature on socio-economic status and mortality is marked by a persisting absence of Germany.<sup>1</sup> This scarcity of studies is probably owed to the lack of large high quality datasets. Luckily, the situation has changed since the inception of the Research Data Centre of the German Pension Insurance (Forschungsdatenzentrum der Rentenversicherung). Using these data enables us to make a number of, albeit small, contributions to the existing literature.

Most obviously, we document mortality inequalities among elderly men in Germany. The data permits us to compare the regions of the former GDR with the rest of Germany. Due to the very different institutions for forty years, differences may well be expected. In addition, the German pension system enables us to measure socio-economic status by means of a variable that we term lifetime earnings. It is a discounted sum of pensionable earnings over the life-cycle. We argue that this is a very broad measure of socio-economic status that is also readily usable for the retired population. Finally, due to the large size of our dataset, we do not need to recur to any parametric assumptions on the structure of the relationship between lifetime earnings and mortality. Hence we are able to provide credible estimates of life expectancies which provide a summary measure of mortality that is readily understood in terms of meaning and magnitude, which is not the case for many other typically used measures.

For the remaining life expectancy at age 65, our results indicate a lower bound of six

---

<sup>1</sup>See for example Mackenbach et al. (2003) or Huisman et al. (2004) and the references cited therein.

years on the difference between the lowest and the highest earnings group considered in our study. That is a difference of almost fifty percent measured from the lowest group. Between the two extremes, life expectancy rises almost linearly in lifetime earnings. Despite the fact that we do find lower overall mortality in the West than in the East, our results show similar life expectancies within income groups. The unconditional difference hence comes from composition effects.

Our results document a pure correlation between lifetime earnings and mortality. From our estimates, nothing can be said about the underlying pathways that lead to these figures. In general, three broad channels of causality can be imagined. Epidemiologists stress the importance of causality from income to health (Marmot 1999). Economists are often preoccupied with quantifying the reverse direction by which it is a low health status that impairs current and future earnings capacity (Smith 2004). A third explanation is that there are one or more underlying factors determining both income and health. Among many potential candidates are genetics, ability, intelligence social skills, networks, and other background or early life factors such as parental income (Case, Lubotsky, and Paxson 2002), or education (Lleras-Muney 2005). We see our contribution in documenting a strong relationship between lifetime earnings and life expectancy in Germany that calls for more research on its origins.

The structure of this paper is the following. We first describe the data that we use in Section 2, with particular emphasis on the German pension system and the calculation of the central explanatory variable. Section 3 contains the presentation of our results in three stages before turning to some international comparisons. Finally, section 4 concludes.

## 2 Data

We use a very large dataset of administrative records from the German Public Pension System. For reasons explained in section 2.2, we consider only male individuals. Our data cover more than eighty percent of the whole male population born 1936 and earlier. Consider the first and fifth column of Table 1 in the Appendix on page 30. These show our sample sizes and what part of the general population is covered by our data broken down by age group and place of residence. Coverage is about three quarters in the West, but it drops substantially at older ages to less than two thirds. Part of this may be explained by overestimation of the general population in the official public records (Jdanov, Scholz, and Shkolnikov 2005), although we use numbers from the Human Mortality Database that already contain a correction for this issue. Another part is likely to be differential mortality - people outside the system tend to have a higher socio-economic status. In terms of selection bias, we do not see any reason why results on differential mortality among the pensioners covered by our data should not extend qualitatively to the rest of the population. Numbers in the East are much nicer with coverages of up to 98 percent. This is because in the former GDR, virtually everybody was insured within the state pension system.

In the next section we turn to the derivation of the central variable of our analysis, an internal measure of the Public Pension System that is used to calculate pension benefits. It is called personal earnings points and serves reasonably well as an indicator of total lifetime earnings. Section 2.2 presents descriptive statistics on the variables used in our analysis and deals with the creation of our dataset from several sources of the administrative records.

## 2.1 The German Public Pension System

The German Public Pension System in the form that is relevant for the cohorts studied in this analysis is a pay-as-you-go system based on a single tier. Benefits are directly related to personal earnings over the life-cycle.<sup>2</sup> This section provides a very brief introduction to those parts of the system that are relevant for the purposes of the paper. Our description is based on Börsch-Supan and Wilke (2004) and VDR (2004). The system covers all private and public sector employees, excluding only civil servants (about 7 percent of the workforce) and most self-employed (about 9 percent). The latter can self-insure in the system, we will get back to this in section 2.2. Our focus is on old-age pensions which are paid to all retirees age 65 and above. By the end of the calendar year in the course of which age 65 is completed, virtually everybody is retired.<sup>3</sup>

Key to the system are the so-called earnings points, which essentially are a measure of the relative annual earnings position. In any given year  $t$ , the earnings points for contribution periods ( $EP^{CP}$ ) of an individual  $i$  are calculated as:

$$(1) \quad EP_{it}^{CP} = \frac{\text{pensionable earnings}_{it}}{\text{average pensionable earnings}_t}$$

In 2002, pensionable earnings were the first 4,500 Euro of gross monthly wages if the individual's earnings were above the minimum earnings threshold of 325 Euro. A subset of our data contains the sum of  $EP_{it}^{CP}$  over all  $t$  with relevant contributions for each individual

---

<sup>2</sup>For a simple taxonomy of pension systems and an international comparison cf. OECD (2005)

<sup>3</sup>For the 1936 cohort (the youngest included in our analysis), internal statistics of the Deutsche Rentenversicherung show that only 0.56 percent were not retired on 1st January 2002, the starting point of our analysis.

(we call this variable  $EP_i^{CP}$ ). Note that this variable is subject to a difficult form of right censoring because of the annual upper limit to pensionable income. Hence we know only a lower bound for the earnings of people with high  $EP_i^{CP}$ . This has to be kept in mind when interpreting the results. We note that because of the division by average pensionable earnings in (1), the discount rate inherent to  $EP_i^{CP}$  is the annual wage growth rate.

For administrative reasons,  $EP_i^{CP}$  is available only for individuals who retired after 1992, the year of a major reform to the system. The measure present for all persons in our dataset is called personal earnings points ( $EP^{pers}$ ). These are calculated as follows:

$$(2) \quad EP_i^{pers} = (EP_i^{CP} + EP_i^{NCP}) \cdot AF_i$$

$EP^{NCP}$  stands for earnings points from non-contributory periods. These stem from spells with no contributions at all which are nonetheless relevant for some pension benefits. These include, for example, unemployment spells, some allowance for advanced education, and the like. The adjustment factor  $AF_i$  scales down benefits in the case of early retirement after the 1992 reform. For our purposes it also serves to capture a type of minimum pension benefit for low earnings (Mindestentgeltpunkte bei geringem Arbeitsentgelt) that took place before 1992.

Individual pension payments are obtained directly from  $EP_i^{pers}$  by multiplication with the current pension value that is common for all pensioners. In 2002 it was 25.86 Euro for  $EP_i^{pers}$  earned in western Germany and 22.97 Euro for those  $EP_i^{pers}$  earned in the GDR. Hence some 50  $EP^{pers}$  translate into a monthly gross pension payment of 1293.00 Euro (1148.50 Euro) in the West (East). The current pension value is adjusted annually accord-

ing to complex procedures, this does not impact upon our analysis. We only need the fact that  $EP^{CP}$  and  $EP^{pers}$  remain constant once an individual receives an old-age pension. For  $EP^{pers}$  there are some minor qualifications to this, for example due to divorce or moving abroad. Since we only include pensioners living in Germany (see section 2.2), the latter does not impact upon our analysis and we treat the former as negligible.<sup>4</sup>

Obviously we prefer  $EP^{CP}$  as a measure of lifetime earnings. For cohorts born after 1928 we can compare both measures. Correlations are very high with  $\rho \simeq 0.95$ . We present the results of a comparative analysis of mortality experiences based on both different measures in section 3.2. Our results show that the distinction is not all too important in terms of describing the mortality experience by earnings group for ages 65 to 73. For the calculation of life expectancies we also need the mortality experience of older cohorts. Hence we extrapolate the similarity result and interpret  $EP^{pers}$  as lifetime earnings although there is larger error inherent to it. The nice feature of these variable is that they give us a measure of long-run earnings. This is a much nicer measure for socio-economic status than current income typically recorded in surveys. The latter is often blurred by transitory fluctuations which are surprisingly high at first glance. These may lead to serious biases as documented in Haider and Solon (forthcoming). Our discounted sum of lifetime earnings misses out on some things typically included in the income definition (for example bequests, capital income or transfers). Bearing the incompleteness in mind, we use lifetime earnings and lifetime income as synonyms in the remainder of the paper.

---

<sup>4</sup>Calculations based on the “Versorgungsausgleichstatistik” show that changes in  $EP^{pers}$  due to divorce affect only 2.6% of the cases in our sample.



Let us close this section with a quick illustration of the monthly pensionable income necessary to accumulate a certain amount of earnings points. In 2002, the monthly gross wage that yielded one  $EP^{pers}$  was about 2,400 Euro. For simplicity, assume that this number remains constant over an individual's working life. Hence, to accumulate 50  $EP^{pers}$ , a person with this wage would have to work 50 years. If average earnings over the life-cycle were 3,000 Euro, 40 full years of contribution would be sufficient to accumulate the same amount of  $EP^{pers}$ .

## 2.2 Description of the Dataset

The administration of the German public pension system is marked by a variety of statutory bodies. Traditionally, there have been regional pension insurance institutes for workers, a federal institute for salaried employees, and three profession-specific institutes. Except for miners, legal regulations have been the same since 1949, however. All pension insurance institutes are required by law to report statistics of all pensioners as of the end of each year as well as statistics of those pensioners who died during that year to their umbrella association.<sup>5</sup> We have access to this data.<sup>6</sup> Because only selective characteristics of the

---

<sup>5</sup>This used to be the Verband Deutscher Rentenversicherungsträger (VDR). After a major organisational reform that took effect on 1st October 2005, its duties are fulfilled by the Deutsche Rentenversicherung Bund, the federal pensions institute.

<sup>6</sup>Traditionally, only aggregate statistics were published. This has changed since the beginning of 2004 with the creation of the Research Data Centre of the Public Pension Insurance. Information about datasets and access procedures can be found at <http://www.fdz-rv.de>. The process of data collection is described in Rehfeld (2001)

original administrative records enter our dataset, some important remarks about structure and peculiarities of the data are in order.

For one thing, there is no way to link members of the same household. Ideally, we would want to use lifetime household income as a relevant measure to correlate with mortality. Because of the low female labour force participation in the cohorts relevant to our analysis, we exclude them from the analysis. It is simply unclear what the household income position of women with low  $EP^{CP}$  is because of the dominance of male earnings in total household income.

For all individuals in our dataset, we have the year and month of their birth available. We only use individuals born in 1936 and earlier because of possible health and income differentials in early retirement. Put differently, if we were to use younger pensioners, most likely we would not have a random sample available. A further demographic variable is the place of residence in three categories (eastern Germany, western Germany, foreign). We exclude people with foreign domicile (2.3%) in order to work with a subset of those recorded in official population statistics. Including them did not cause any visible differences in the mortality estimates. The data contain all deaths in 2002, these are recorded on a monthly level. The Appendix contains descriptive statistics for the entire sample in Table 2. Those restricted to pensioners born after 1928 are listed in Table 3.

In terms of variables related to pension payments, the two most important ones are those described in the last section. Note that  $EP^{pers}$  is on average about 3.1 points (6.6 percent) higher than  $EP^{CP}$  in columns 1 and 5 of Table 3. We also have the corresponding variables with respect to the length of being insured. These are pension-relevant insurance periods

( $IP^{PR}$ ) and contribution periods (CP). The former are comprised of the latter and non-contributory periods eligible for pension benefits (for examples see Section 2.1). Last, we use information on the type of health insurance coverage. Employees are mandatorily covered in the public mutual funds system up to an earnings threshold that was 75% of the maximum pensionable earnings until 2003. Individuals above that threshold, the self-employed and civil servants can either insure voluntarily in the system or opt out to join a private insurance company. A small subgroup of pensioners in our sample is insured under foreign law, these persons usually worked in Germany only for short periods of time. The arrangement of the last employment spell usually carries over to retirement. We can identify the three groups (mandatorily insured / voluntarily or privately insured / insured under foreign law) in our data.

The reason why this becomes important lies in work biographies that are not confined to a single system of pension insurance. As an example, take somebody who is employed for ten years and then becomes a civil servant for the rest of his working life. If we used  $EP^{CP}$  as a measure of his lifetime earnings, we would make a huge error because he had large earnings outside the system. The health insurance variable as well as the length of pension insurance periods enable us to (partially) control for such cases. This is why we estimate life expectancies for four different subgroups of the population: All pensioners (“All”), individuals mandatorily insured within the public health insurance scheme (“HI”), those with more than 25 years of pension-relevant insurance periods (“25Y”), both the “HI” and “25Y” restrictions imposed (“HI25Y”). These make up the remaining columns of Tables 1 to 3. Looking at the population coverage in Table 1, it becomes clear that it drops to two

thirds in the West if both restrictions are active. The decline is particularly sharp at older ages because the differential mortality effect is reinforced (see below in section 3.1) and because there are more missing values at old ages. The only variable that is completely available is  $EP^{pers}$  because it is central for the pension payment. In conversations with statisticians at the Deutsche Rentenversicherung Bund, we tried to evaluate the influence of systematic effects on missing values. Except for the cases that we mention, there is no reason to expect a missing at random assumption to be violated. In the East the picture is much nicer again with coverage rates above ninety percent except for the very old ages.

There are three more variables available that we do not consider in the presentation of differential mortality by lifetime earnings for clarity reasons. These are citizenship in two categories (German/non-German), whether a pension entitlement for repatriates forms part of the total pension, and whether  $EP^{pers}$  includes a scaling-up of raw earnings points because of low earnings before 1992 (cf. Section 2.1). Confining our analysis to Germans that do not fall into either of the last two categories did not substantially alter any results, tables and graphs are available from the authors upon request.

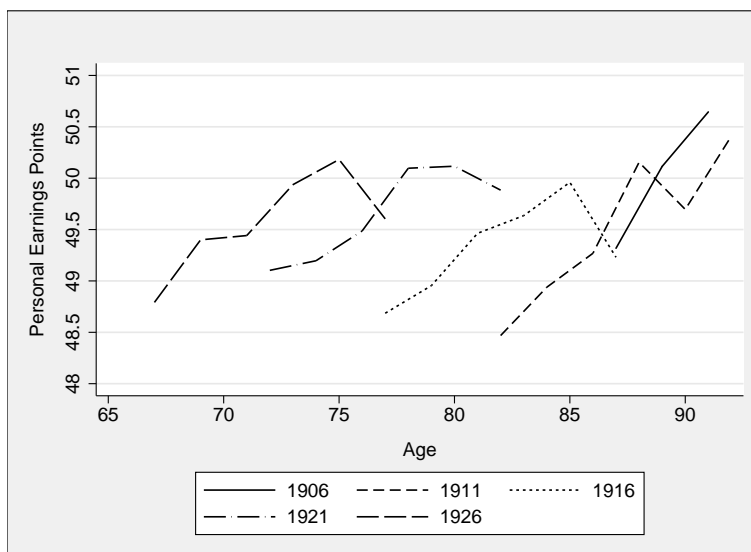
### 3 Results

We start with a brief and simple motivation of our analysis. Figure 1 shows the average number of personal earnings points by birth cohort.<sup>7</sup> The general pattern is very clear - they rise over time. Remember from section 2.1 that  $EP^{pers}$  does not change anymore within

---

<sup>7</sup>These are based on 1%-samples of all pensioners in 1993, 1995, 1997, 1999, and 2001 (DRV 2005a)

Figure 1:  $EP^{pers}$  by Cohort, 1993-2001



*Note:* All pensioners with at least 20  $EP^{pers}$  are included (selection All). The data are collected biannually, cohorts are based on five-year birth intervals. *Source:* Forschungsdatenzentrum der Rentenversicherung, SUFRTNWFXXXXDemo, SUFRTNBNXXXXDemo, own calculations.

persons after entering retirement. Hence, the only reason why it can change over time is by composition of the cohort which in turn is almost entirely determined by mortality. The rise is then due to persons with less than average  $EP^{pers}$  dying relatively more frequently than those with higher earnings points.

Shedding more light on this relationship by means of period life tables is the purpose of this chapter. Ideally, we would prefer a cohort analysis in the spirit of our motivational graphs. Given the structure of the administrative records, however, there is no straightforward way to follow cohorts over time and we confine our analysis to all cohorts in 2002. We present our results in three parts. First, we look at the mortality experience of all German pensioners by means of remaining life expectancy at age 65 ( $e_{65}$ ) and annual mortality rates. We compare the two earnings points measures in the second part of the chapter.

Because of data availability, the best we can do is to consider the probability of reaching age 74, conditional on reaching age 65. Finally, we contrast the mortality experience of persons living in the former GDR with those residing in the western part of Germany.

Throughout the analysis, we divide the sample into eleven equally spaced groups of earnings points. We then present statistics for each of these. Due to the extremely large sample size, we can afford the luxury of not recurring to any parametric assumptions in calculating mortality rates. Lifetables are based on the classic Chiang (1984) formulas, confidence intervals are calculated via a bootstrap procedure with 1000 replications.<sup>8</sup>

### 3.1 Remaining Life Expectancy at Age 65 by $EP^{pers}$

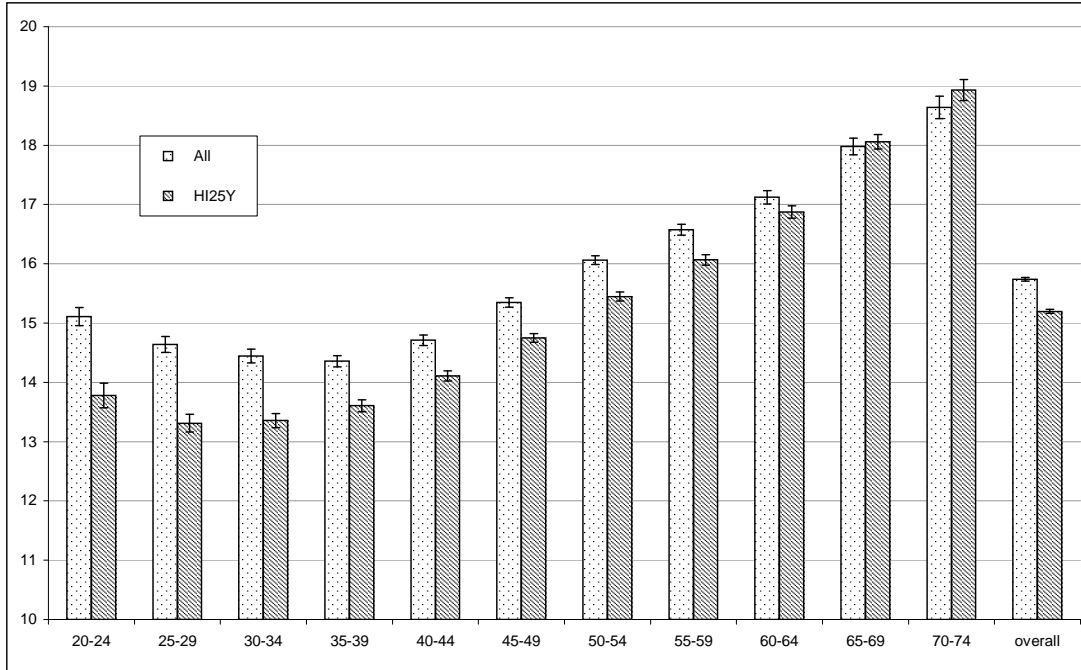
First consider the light bars in Figure 2. These depict  $e_{65}$  for the full sample of pensioners that we use. Overall remaining life expectancy is at 15.74 years, we postpone a comparison to estimates for the general population until the end of this section. The mortality estimates by earnings points group range from 14.35 years (35-39  $EP^{pers}$ ) to 18.65 years (70-74  $EP^{pers}$ ). Between these two extremes, life expectancy rises roughly linearly over the groups. All differences among them are statistically significant at any typical confidence level.

The most striking finding at first glance is that the minimum life expectancy is reached close to the middle of the table and not in the lowest income group. At the bottom of the distribution,  $e_{65}$  is up to more than fifteen years again. This is quite contrary to overwhelming international evidence which indicates a monotone and positive relationship between income and life expectancy. However, there is a plausible reason at hand. As

---

<sup>8</sup>We are very grateful to Evgueni Andreev for sharing his VBA code for bootstrapping life-tables.

Figure 2: Remaining Life Expectancy at Age 65 in years by  $EP^{pers}$



*Note:* Comparison of all pensioners with the respective amount of  $EP^{pers}$  and those who are mandatorily insured in the public health insurance scheme with at least 25 years of pension-relevant insurance periods (HI25Y). The vertical bars indicate 95% confidence intervals

explained in the last section, we expect a very heterogeneous group at the lower end of the distribution because of persons who were covered by the pension system only during parts of their working life. These are typically well-earning academics who would be at the right tail of the distribution if we were to observe their full earnings history. To take a colourful example, we would expect to find production line workers next to their company's CEO in these groups.

A way to shed light on this issue is to exclude those for whom lifetime earnings are not observed very well. We try to do so by selecting only those who are mandatorily enrolled in the public mutual funds health insurance system or those who spent at least 25 years in

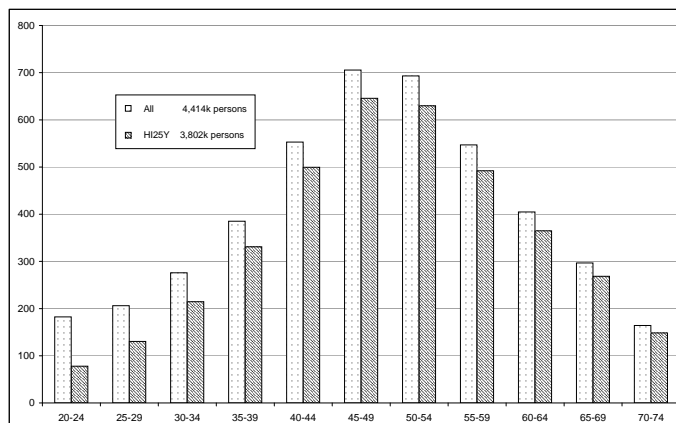
the system. The dark bars in Figure 2 indicate the results if we impose both restrictions. For brevity reasons we do not present the results if selection is based upon one criterion only. Again the tables are available upon request.

To begin with, note that overall remaining life expectancy drops substantially by more than half a year. Hence the fourteen percent that we excluded from the original sample must have a much higher life expectancy than the remaining selection. While it rises slightly in the top two earnings points categories, it drops significantly in all other subgroups. The decline is particularly pronounced in the lowest income categories. Minimum  $e_{65}$  is now 13.31 years for those with 25-29  $EP^{pers}$ , the slight rise for the lowest income category is only borderline significant. We suspect this differential drop to be a combination of two effects. On the one hand, the relative size of the sample that is excluded is much higher at lower earnings points levels (57% in the lowest category as compared to less than 10% in the top seven classes, see Figure 3). On the other hand, if the excluded group is relatively homogeneous, the differential in  $e_{65}$  is largest in the lower categories under the hypothesis of a monotonic relationship between income and mortality.

We presume that this control is far from being perfect, hence we only have a lower bound for the mortality differential. This argument is reinforced through the right censoring of annual  $EP_{it}^{pers}$  - neither can we clearly identify the top nor the bottom earners. Note that somebody in the lowest mortality group of 30-34  $EP^{pers}$  still would have worked 40 years at 3/4 of average wage. If somebody was living on social assistance most of the time, he would not even enter our dataset. However, even this lower bound on differences in life expectancy is quite substantial. Taking the results from the “HI25Y” selection and the



Figure 3: The distribution of  $EP^{pers}$  by Sample Selection

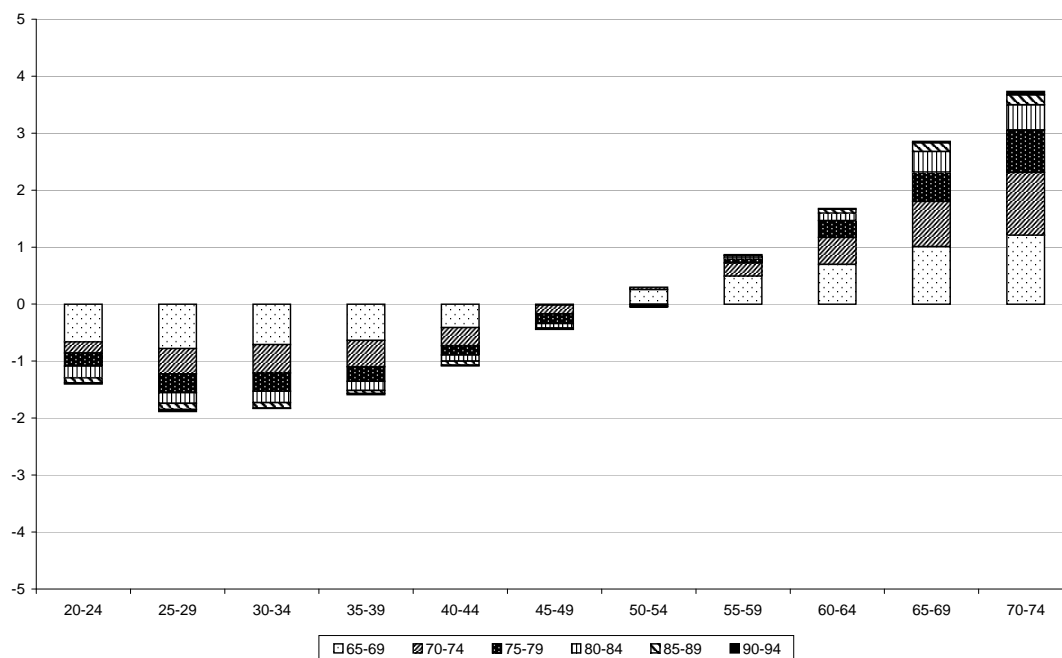


*Note:* “All” includes all available observations, “HI25Y” incorporates only those mandatorily enrolled in the public health insurance scheme with at least 25 years of pension-relevant insurance periods.

unconditional  $e_{65}$  as a starting point, persons in the highest income group can expect to live 25 percent longer. On the other hand, if only 25-29  $EP^{pers}$  were accumulated, it is 12 percent less.

Our results compare quite well with all-population mortality. Official statistics indicate a remaining life expectancy at age 65 for German males in 2002 of 16.08 years (HMD 2006) This is about 4 months higher than our estimates for the full sample indicate. In terms of socio-economic status and mortality experience, most of the persons not covered in our data should be roughly comparable to those in that were excluded when we imposed the “HI25Y” restriction. Simplifying the matter a bit, the main difference between them is that one group worked for a few years in a job that covered them in the public pension system and then changed to another one; the other group started in such another job already. Following this argument we did expect a qualitatively similar rise in life expectancy if we move from “All” to the full population as the one that we see when moving from “HI25Y”

Figure 4: Decomposition of Life Expectancy Differentials by Age Group, relative to Unconditional Mortality

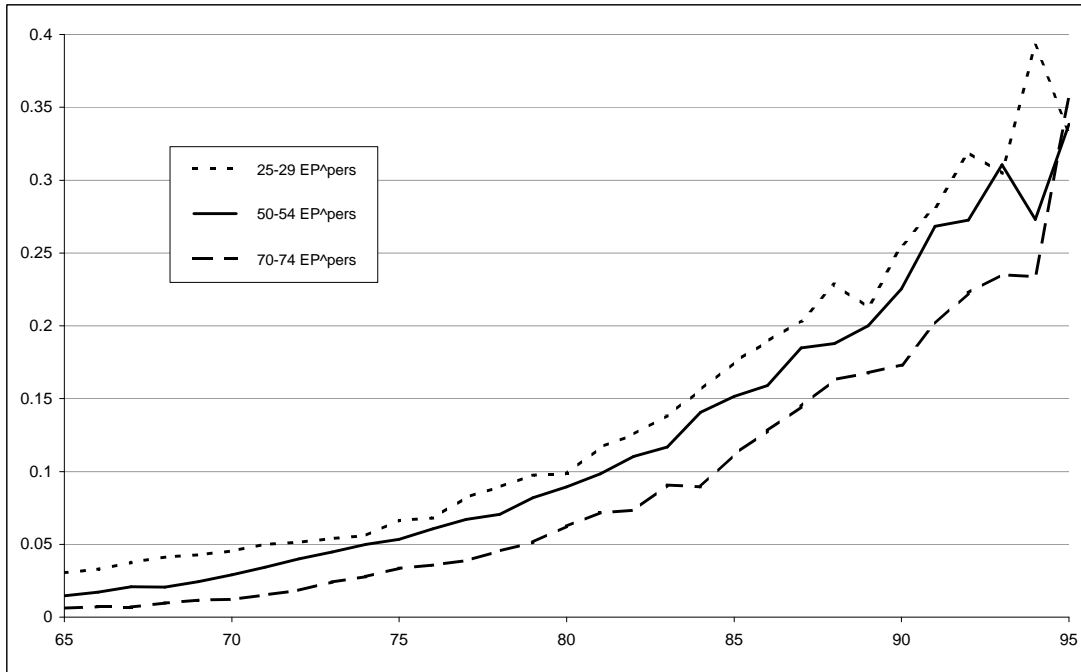


*Note:* Only the persons mandatorily enrolled in the public health insurance scheme with at least 25 years of pension-relevant insurance periods are included in the analysis (selection HI25Y).

to “All”.

Using the method due to Arriaga (1984), we can decompose the contributions to life expectancy by age group. Figure 4 breaks down the mortality differentials for each  $EP^{pers}$ -group by 5-year age bands relative to unconditional mortality. Basically, this works by replacing the respective age classes of the income-differentiated death probabilities by the unconditional ones. It becomes immediately clear that the bulk of the difference between income groups stems from mortality before age 75. Beyond 80 years of age, non-negligible discrepancies exist only for the two highest earnings classes. Given the high weight that  $e_{65}$  puts on the lower age classes, this result does not come as a surprise to us.

Figure 5: Mortality Rate by  $EP^{pers}$



*Note:* Only the persons mandatorily enrolled in the public health insurance scheme with at least 25 years of pension-relevant insurance periods are included in the analysis (selection HI25Y).

Finally, Figure 5 shows mortality rates for selected income groups under the “HI25Y” restriction. They were calculated separately for each one-year age band. The most salient feature of the graph is that their shape is very similar, they seem to differ by little more than a parallel shift. Confidence bands for each age are not shown in order to keep the graphs readable. Until age 74, all three are statistically different from each other at the 95%-level. Mortality rates in the highest income group are significantly lower than the other two even until age 88.

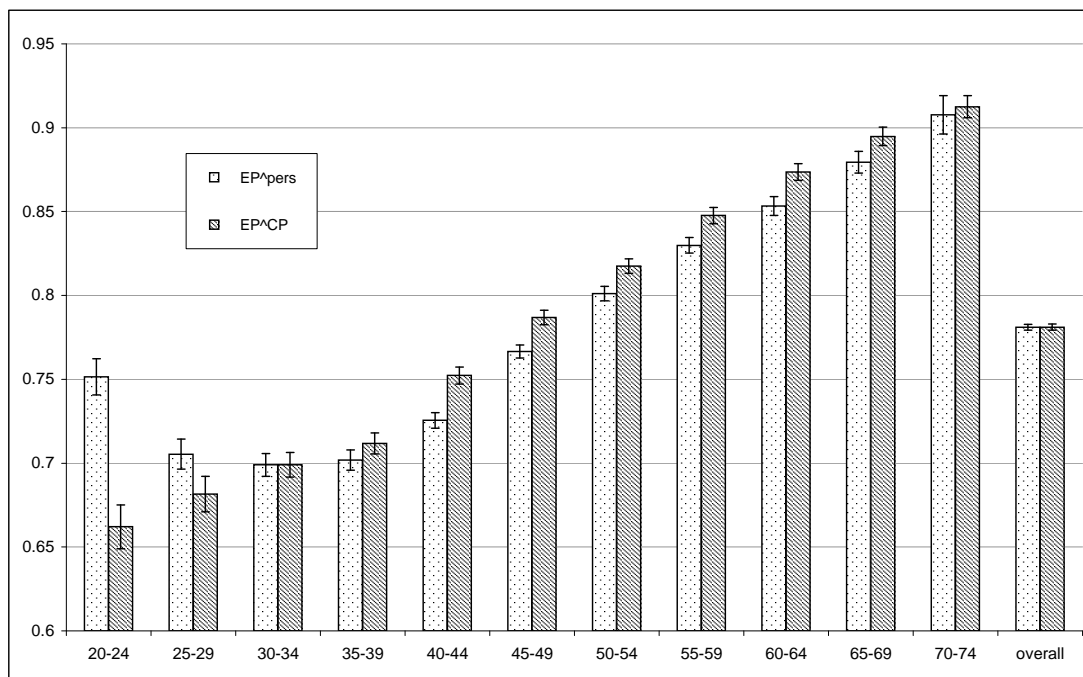
### 3.2 Mortality by $EP^{pers}$ and $EP^{CP}$

The purpose of this section is to contrast the two different measures of lifetime earnings for those persons where both are available. These are the cohorts born 1929 and later because most of them did not retire before 1992. Remaining life expectancy is not a suitable summary statistic anymore because we do not have any information on old-age mortality conditional on  $EP^{CP}$ . As an alternative, we chose the probability of reaching age 74 (this is the highest we can do) conditional on reaching age 65 ( $\mathbb{P}_{65}\{74\}$ ). Results are shown in Figure 6. The light bars depict  $\mathbb{P}_{65}\{74\}$  conditional on  $EP^{pers}$ , the dark ones show the corresponding values based on  $EP^{CP}$ .

Overall probabilities are identical at 76.4 percent because we use the same sample in both cases. In the case of  $EP^{pers}$ , we find very much the same pattern as for life expectancy in the last section. There is a linear decline from the highest income class to those persons with 35-39  $EP^{pers}$  which then levels out and rises again at the very bottom. The dark bars look quite similar, but there are some important differences. On the one hand, probabilities are slightly higher for all groups with more than 35  $EP^{CP}$ . However, the decline does not level out at this point but continues linearly to the very lowest group. Based on this measure, 65-year-old individuals in the highest earnings class have a ninety percent probability of reaching age 74. Less than two thirds in the bottom category survive to this age.

Our conclusions from this exercise are twofold. On the one hand, it does not matter much whether one uses  $EP^{pers}$  or  $EP^{CP}$  if one is interested in the mortality experience for

Figure 6: Probability of Reaching Age 74 at Age 65 by  $EP^{pers}$



*Note:* Only the persons mandatorily enrolled in the public health insurance scheme with at least 25 years of pension-relevant insurance periods are included in the analysis (selection HI25Y). The vertical bars indicate 95% confidence intervals.

individuals with at least 35  $EP^{pers}$ . Our summary statistics differ by not much more than a constant that is mostly explained by the average differential of 3 points between the two measures (see the first two rows of Table 3). On the other hand, the choice of variable does matter in lower categories. By using  $EP^{CP}$  as a measure of lifetime earnings, we can reproduce the monotonic relationship to mortality that is documented in international studies.

In the course of the analysis of this chapter we also checked whether differences exist if we condition on actual contribution periods (CP) rather than pension-relevant insurance periods ( $IP^{PR}$ ). One might have expected analogous differences as in the case of the two earnings measures. However, none such differences became apparent which is why we do not include a graph on that point.

### 3.3 Life Expectancy at Age 65 by $EP^{pers}$ and Place of Residence

In this section, we compare the mortality experience of people living in the former socialist part of Germany with those living in western Germany. This is particularly interesting to look at because of the very different biographies of people living in either part.<sup>9</sup> One could expect several things to happen in the eastern part. There could be a long-run effect of the more equal distribution of socio-economic status during the socialist era, resulting in smaller mortality differentials. The opposite story might read that the sharp

---

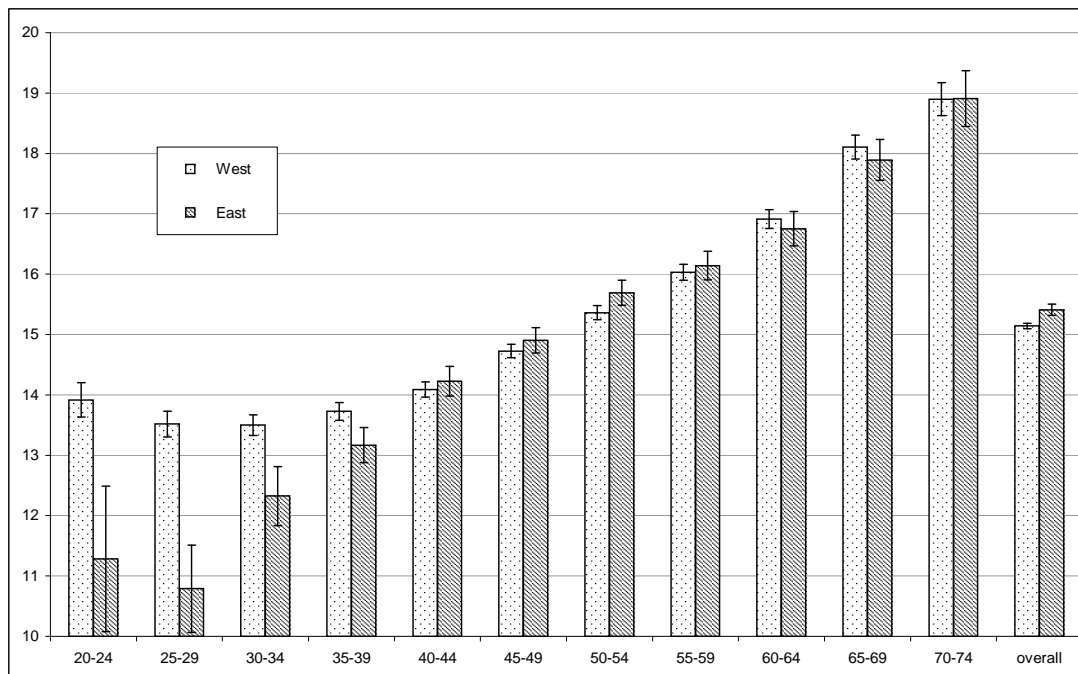
<sup>9</sup>In our interpretations of this section, we neglect migration between the two regions. People in our sample were at least 53 years old at the time of reunification, so we do not expect any substantial migration movements that would blur our results

transformation in the early 1990s led to higher inequality than in the West. Finally, one could imagine relatively quick adaptation to the new institutional arrangement, hence a picture that parallels that in the West.

Naturally, the first thing to evaluate is  $e_{65}$  not stratified by income. We find it to be 15.83 years in the West and 15.41 years in the East. The difference is statistically significant at any common confidence level. This compares to full-population estimates from official statistics of 16.19 years and 15.39 years, respectively. In the West we have the effect described above: The 23% of the population not included in our sample tend to have a lower mortality than the pensioners. On the other hand, the 94% coverage in the East leads to almost identical estimates of all-population mortality, so we are very confident to have a complete depiction of the full population there. Our findings are consistent with the converging mortality experiences in the East and in the West that have been documented by several authors (cf. for example Nolte, Shkolnikov, and McKee (2000)).

Again, the main analysis concerns the comparison of life expectancy by income group. As in the previous sections, we select upon mandatory enrollment in the public mutual funds health care system and 25 years of coverage in the pension system. The reason not to consider all-pensioner mortality here is that groups are more comparable in the restricted sample. As discussed in Section 2.2, we expect much more heterogeneity in terms of earnings outside the system in the West than in the East, particularly in the lowest categories. This is because in the former GDR, virtually everybody was insured in the system. Hence we would obtain a stronger bias within income classes if we did not impose the restrictions. The reason is the larger heterogeneity with respect to socio-

Figure 7: Remaining Life Expectancy at Age 65 in years by Place of Residence and  $EP^{pers}$



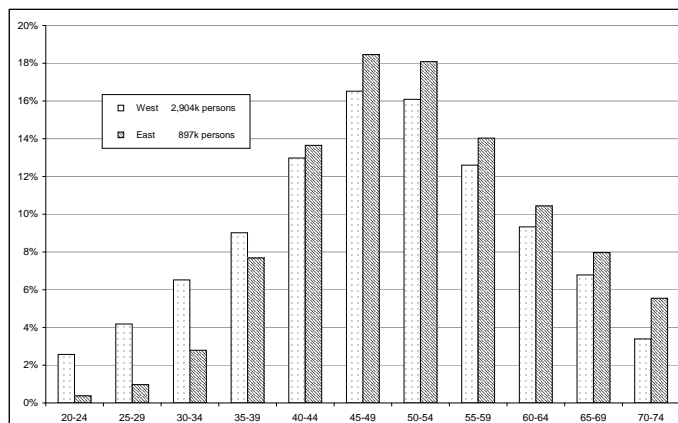
*Note:* Only the persons mandatorily insured in the public health insurance scheme with at least 25 years of pension-relevant insurance periods are included in the analysis (selection HI25Y). The vertical bars indicate 95% confidence intervals.

economic status in the West. The restriction of the sample makes the analysis within categories more meaningful, but a comparison of the unconditional life expectancies does not make much sense. This becomes clear from the rightmost bars in Figure 7. While  $e_{65}$  does not change for pensioners in the east of Germany, imposing the restriction leads to a drop in remaining life expectancy to 15.14 years in the West. This reversal is due to the fact that we excluded a large group of men with high socio-economic status in the West and only few persons in the East.

Comparing the graphs for both regions in Figure 7, it becomes apparent that no difference between them can be asserted for groups with more than 40  $EP^{pers}$ . In all groups



Figure 8: The distribution of  $EP^{pers}$  by Place of Residence



*Note:* Only the persons mandatorily insured in the public health insurance scheme with at least 25 years of pension-relevant insurance periods are included in the analysis (selection HI25Y).

below that level, life expectancy in the West emerges larger than in the East. Most notably, we find a close to strongly monotone relationship between income and life expectancy in the East (the rise in the bottom group is not detectable in a statistical sense for either region). In the lowest two groups in the East,  $e_{65}$  is only around eleven years, eight years less than at the top of the distribution. We do not take this as evidence for more inequality in the East. It is more likely that our selection procedure still does not enable us to pinpoint people with low socio-economic status in the West, just as the analysis from Section 3.2 suggests. The higher life expectancy in the East for this particular selection arises because the distribution of earnings points is shifted to the right as compared to western Germany, see Figure 8. In other words, there are relatively more people in the higher income classes with a longer life expectancy.

A second thing to note is the striking similarity of the results in the income classes from 40  $EP^{pers}$  on upwards. This is somewhat surprising given that biographies have been quite

different. One part lived for forty years in a socialist country with severely restricted civil rights, comparatively small wage differentials, high job security, an economy with a large amount of goods rationing, etc. People on the other side of the wall experienced much higher wage differentials, somewhat lower job security, a market economy, and quite well protected civil rights. Only from 1990 on - the cohorts in our analysis were already at least 53 years old - institutional arrangements started to converge, although large differences remain, for example on labour markets. Furthermore, not even our income measure is strictly comparable. If  $EP^{pers}$  is interpreted as actual retirement income, there is still a 13%-difference between the two (see Section 2.1) that is not compensated by lower prices in the East. The interpretation of lifetime earnings remains difficult because actual purchasing power from one  $EP^{pers}$  was extremely different in the GDR as compared to the FRG. Our understanding of the similarity finding is that  $EP^{pers}$  contains many “deep” parameters that are related to mortality and survive institutional arrangements. Examples include education, intelligence, social skills and networks, or genetics. Put differently,  $EP^{pers}$  could be a very broad measure of socio-economic status that yields a ranking of people which is remarkably stable across some forms of societal organisation. This argument is reinforced through institutional features from the GDR. Higher qualified employees were more likely to have a so-called “Voluntary Additional Insurance”, resulting in higher differentiation of earnings points than that of current labour income. Hence,  $EP^{pers}$  would reflect a broader measure of socio-economic status than the level of material resources.

Our analysis sheds a new light on previous findings from the epidemiological literature. Several authors reported a larger predictive power of income for health measures in the

West than in the East (Mielck et al. (2000), Nolte and McKee (2004)). Our results suggest this to be a consequence of measurement error, unless the link between morbidity and mortality works differently between both parts of Germany. The relevant income concept as a marker for socio-economic status is a long-run measure, the surveys employed in the cited studies contain a current income variable. Transitory fluctuations, for example due to high unemployment in the East, may well impact strongly upon the analysis.

### 3.4 International Comparisons

In this section, we place our results in the context of the literature on differential mortality. Closest in terms of regional proximity and research question is certainly Reil-Held (2000). She uses survey panel data and finds a life expectancy differential of 10 years between the top and the bottom quartile of the income distribution, employing a somewhat broader household income variable that is averaged across time. This is in the bivariate analysis that is comparable to our approach, confidence intervals are not reported. Both results go together very well - she computes  $e_{37}$  and the four-year difference between our findings will be more than accounted for by mortality until age 65. We take this as a supporting evidence for our result of a lower bound of six years on the difference in  $e_{65}$  between the top and the bottom income class.

Expanding the region under consideration and the variables measuring socio-economic status, the study by Huisman et al. (2004) serves well to place our results in a European context. They analyse differential mortality by education and housing tenure and find

results that are quite similar to ours in qualitative terms. Throughout the eleven countries included in their study, relative mortality risks declined in age, but absolute mortality differences remained constant or increased until old age. This is exactly what we find from Figure 5. It is impossible to compare numbers because of the differences in covariates, furthermore the authors do not report life expectancies.

Further corroborative evidence for our findings comes from the US - Deaton and Paxson (2004) report  $e_{25}$  to be about 10 years lower for members of families with an annual income of less than \$5,000 as compared to those earning more than \$50,000. Findings from Attanasio and Hoynes (2000) suggest that the gradient linking income and mortality is steepest at the bottom of the distribution. This is supportive of our view that we cannot identify people at the lower end of the distribution very well. Finally, turning back to Europe, Attanasio and Emmerson (2003) show that the position in the wealth ranking has a large impact on survival probabilities in the UK. Palme and Sandgren (2004) are able to construct a measure of lifetime income for a cohort born in 1928 in the city of Malmö, Sweden. They find this variable to be only marginally significant in the case of prime-age mortality. However, their sample size is rather small and they conduct their analyses conditional on parental income. Overall, we conclude that our findings for Germany fit in very well with the existing literature on socio-economic status and mortality.

## 4 Conclusions

We found large differentials in remaining life expectancy at age 65 across classes of lifetime earnings in Germany. Due to the nature of our data, we were only able to put a lower bound of six years on the difference from the lowest to the highest income group. Between the two extremes, life expectancy rises linearly in earnings. Since we employed entirely nonparametric methods, this is not an artefact of any assumed functional form.

Comparing the former GDR with the rest of Germany, we found virtually no differences in life expectancy within income groups, except for the lowest groups. However, the latter seems to be due to peculiarities of our data. Hence, any aggregate difference in life expectancy is likely owed to composition effects. This finding is quite remarkable because of the very different institutional design people in either part were faced with during the prime ages of their working life.

What do we learn from these results? For one thing, there is a very sizeable effect, even among the elderly, waiting to be explained. Which of the three channels mentioned in the introduction is responsible for how much of the differential? Economic analyses suggest that in the socio-economic status to health causation, income is not likely to play a large role (compare, for example, Adams, Hurd, McFadden, Merrill, and Ribeiro (2003), Meer, Miller, and Rosen (2003), or Smith (2004)). Another finding that deserves further illumination is the similarity of life expectancy in the East and in the West conditional on lifetime earnings. Finally, an interesting research question lies in the temporal changes of mortality. How was the gain in life expectancy over the last ten years distributed

over different subgroups of the population? Did the poor or the rich gain more or was it evenly distributed? Is this trend likely to continue? Answers to these questions have a large impact on pension finance, the organisation of nursing home care, and many other important policy questions.

## A Tables

Table 1: Sample Size and Coverage of the General Population by Cohort

Birthyear	West				East			
	All	HI	25Y	HI25Y	All	HI	25Y	HI25Y
before 1902	494 (64%)	396 (51%)	355 (46%)	283 (37%)	154 (96%)	130 (81%)	151 (94%)	127 (79%)
1902 - 1906	7,425 (60%)	6,649 (54%)	5,691 (46%)	5,202 (42%)	2,307 (98%)	2,115 (90%)	2,267 (96%)	2,082 (88%)
1907 - 1911	52,680 (61%)	48,448 (56%)	43,557 (50%)	41,152 (48%)	13,874 (95%)	13,339 (91%)	13,795 (94%)	13,269 (91%)
1912 - 1916	152,514 (71%)	135,688 (63%)	126,505 (59%)	116,739 (54%)	35,139 (96%)	34,281 (93%)	34,987 (95%)	34,158 (93%)
1917 - 1921	321,267 (76%)	288,243 (68%)	268,837 (64%)	247,470 (59%)	73,661 (92%)	72,450 (90%)	73,381 (92%)	72,216 (90%)
1922 - 1926	629,208 (76%)	579,373 (70%)	532,667 (64%)	498,207 (60%)	141,318 (91%)	139,044 (89%)	140,761 (90%)	138,571 (89%)
1927 - 1931	1,015,922 (79%)	931,499 (72%)	927,315 (72%)	855,694 (66%)	272,274 (93%)	265,174 (91%)	271,711 (93%)	264,689 (91%)
1932 - 1936	1,301,622 (78%)	1,155,290 (69%)	1,273,053 (76%)	1,140,150 (68%)	393,716 (97%)	372,431 (92%)	393,438 (97%)	372,273 (92%)
Total	3,481,132 (77%)	3,145,586 (70%)	3,177,980 (70%)	2,904,897 (64%)	932,443 (94%)	898,964 (91%)	930,491 (94%)	897,385 (91%)

*Note:* Sample Size as number of pensioners, coverage of the general population in parentheses. “All” includes all pensioners in the respective age range with more than 20  $EP^{pers}$ . “HI” restricts the sample to individuals who are furthermore mandatorily insured within the public health insurance scheme. “25Y” considers only pensioners with more than 20  $EP^{pers}$  and more than 25 years of pension-relevant insurance periods. “HI25Y” imposes both the “HI” and “25Y” restrictions. Source: DRV (2005b), HMD (2006), own calculations.

Table 2: Descriptive Statistics for all Individuals

Variable	West				East			
	All	HI	25Y	HI25Y	All	HI	25Y	HI25Y
EP <sup>pers</sup>	47.9 (12.8)	48.7 (12.2)	48.7 (12.4)	49.2 (11.9)	52.0 (10.6)	52.2 (10.5)	52.0 (10.6)	52.2 (10.4)
IP <sup>PR</sup>	39.0 (11.4)	39.5 (11.3)	42.3 (4.4)	42.5 (4.1)	44.1 (2.7)	44.2 (2.6)	44.2 (2.0)	44.3 (1.9)
HI=1	0.904	1.000	0.914	1.000	0.964	1.000	0.964	1.000
HI=2	0.092	0.000	0.082	0.000	0.034	0.000	0.034	0.000
HI=3	0.004	0.000	0.004	0.000	0.002	0.000	0.002	0.000
GERMAN	0.968	0.969	0.972	0.973	0.998	0.998	0.998	0.998
REPAT	0.049	0.050	0.052	0.052	0.015	0.015	0.015	0.015
MIN PENS	0.043	0.041	0.046	0.045	0.016	0.015	0.016	0.015
DEAD	0.051	0.053	0.053	0.055	0.051	0.051	0.051	0.051

*Note:* Mean of variables, standard errors in parentheses where appropriate. “HI=1,2,3:” Health insurance coverage by: mandatory public mutual funds, voluntary mutual funds or private insurance company, insurance under foreign law. “GERMAN” shows the fraction of pensioners with German citizenship, “MIN PENS” the fraction entitled for a scaling-up of EP<sup>CP</sup> due to low earnings before 1992, and “DEAD” the fraction that died during 2002.

“All” includes all pensioners in the respective age range with more than 20 EP<sup>pers</sup>. “HI” restricts the sample to individuals who are furthermore mandatorily insured within the public health insurance scheme. “25Y” considers only pensioners with more than 20 EP<sup>pers</sup> and more than 25 years of pension-relevant insurance periods. “HI25Y” imposes both the “HI” and “25Y” restrictions. Source: DRV (2005b), own calculations.



Table 3: Descriptive Statistics for Individuals born after 1928

Variable	West				East			
	All	HI	25Y	HI25Y	All	HI	25Y	HI25Y
EP <sup>pers</sup>	47.6 (12.6)	48.3 (11.9)	48.1 (12.2)	48.6 (11.7)	50.9 (10.3)	51.1 (10.1)	50.9 (10.3)	51.1 (10.1)
EP <sup>CP</sup>	44.9 (12.9)	45.6 (12.3)	45.6 (12.6)	46.0 (12.1)	47.4 (9.8)	47.4 (9.6)	47.4 (9.8)	47.5 (9.6)
IP <sup>PR</sup>	41.0 (8.0)	41.4 (7.8)	42.4 (4.3)	42.6 (4.0)	44.0 (2.4)	44.1 (2.2)	44.1 (2.1)	44.1 (2.0)
CP	37.8 (7.6)	38.3 (7.2)	38.4 (6.9)	38.7 (6.8)	5.9 (6.9)	5.9 (6.8)	5.9 (6.8)	5.9 (6.8)
HI=1	0.896	1.000	0.903	1.000	0.953	1.000	0.954	1.000
HI=2	0.099	0.000	0.092	0.000	0.044	0.000	0.044	0.000
HI=3	0.005	0.000	0.004	0.000	0.002	0.000	0.002	0.000
GERMAN	0.956	0.957	0.962	0.963	0.998	0.998	0.998	0.998
REPAT	0.062	0.063	0.063	0.064	0.018	0.018	0.018	0.018
MIN PENS	0.038	0.036	0.039	0.037	0.020	0.018	0.020	0.018
DEAD	0.028	0.028	0.028	0.029	0.029	0.029	0.029	0.029

*Note:* Mean of variables, standard errors in parentheses where appropriate. “HI=1,2,3:” Health insurance coverage by: mandatory public mutual funds, voluntary mutual funds or private insurance company, insurance under foreign law. “GERMAN” shows the fraction of pensioners with German citizenship, “MIN PENS” the fraction entitled for a scaling-up of EP<sup>CP</sup> due to low earnings before 1992, and “DEAD” the fraction that died during 2002.

“All” includes all pensioners in the respective age range with more than 20 EP<sup>pers</sup>. “HI” restricts the sample to individuals who are furthermore mandatorily insured within the public health insurance scheme. “25Y” considers only pensioners with more than 20 EP<sup>pers</sup> and more than 25 years of pension-relevant insurance periods. “HI25Y” imposes both the “HI” and “25Y” restrictions. Source: DRV (2005b), own calculations.

## References

- ADAMS, H. P., M. D. HURD, D. MCFADDEN, A. MERRILL, AND T. RIBEIRO (2003): “Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status,” *Journal of Econometrics*, 112(1), 3–56.
- ARRIAGA, E. E. (1984): “Measuring and Explaining the Change in Life Expectancy,” *Demography*, 21(1), 83–96.
- ATTANASIO, O., AND C. EMMERSON (2003): “Mortality, Health Status and Wealth,” *Journal of the European Economic Association*, 1(4), 821–850.
- ATTANASIO, O. P., AND H. W. HOYNES (2000): “Differential Mortality and Wealth Accumulation,” *Journal of Human Resources*, 35(1), 1–29.
- BÖRSCH-SUPAN, A., AND C. B. WILKE (2004): “The German Public Pension System: How It Was, How It Will Be,” NBER Working Paper No. 10525.
- CASE, A., D. LUBOTSKY, AND C. PAXSON (2002): “Economic Status and Health in Childhood: The Origins of the Gradient,” *American Economic Review*, 92(5), 1308–1334.
- CHIANG, C. L. (1984): *The Life Table and its Applications*. Krieger Pub. Co., Malabar, Fla.
- DEATON, A., AND C. PAXSON (2004): “Mortality, Education, Income, and Inequality Among American Cohorts,” in *Perspectives on the Economics of Aging*, ed. by D. A. Wise. University of Chicago Press, Chicago, IL.
- DRV (2005a): “Scientific Use File: Themendatensatz Demografie, SUFRT-NWFXXXXDemo, SUFRTNBNXXXXDemo,” Dataset, Deutsche Rentenversicherung Bund.
- (2005b): “Sonderauswertung Fernrechenfile Differentielle Mortalität 2003,” Restricted Use Dataset, Deutsche Rentenversicherung Bund.
- HAIDER, S., AND G. SOLON (forthcoming): “Life-Cycle Variation in the Association between Current and Lifetime Earnings,” *American Economic Review*.
- HMD (2006): “Human Mortality Database,” <http://www.mortality.org>.
- HUISMAN, M., ET AL. (2004): “Socioeconomic Inequalities in Mortality Among Elderly People in 11 European Populations,” *Journal of Epidemiology and Community Health*, 58, 468–475.

- JDANOV, D. A., R. D. SCHOLZ, AND V. M. SHKOLNIKOV (2005): “Official Population Statistics and the Human Mortality Database Estimates of Populations Aged 80+ in Germany and Nine Other European Countries,” *Demographic Research*, 13(14), 335–362.
- LLERAS-MUNEY, A. (2005): “The Relationship Between Education and Adult Mortality in the United States,” *Review of Economic Studies*, 72(1), 189–221.
- MACKENBACH, J. P., ET AL. (2003): “Widening Socioeconomic Inequalities in Mortality in Six Western European Countries,” *International Journal of Epidemiology*, 32, 830–837.
- MARMOT, M. (1999): “Multi-Level Approaches to Understanding Social Determinants,” in *Social Epidemiology*, ed. by L. Berkman, and I. Kawachi. Oxford University Press, Oxford.
- MEER, J., D. L. MILLER, AND H. S. ROSEN (2003): “Exploring the Health-Wealth Nexus,” NBER Working Paper 9554.
- MIELCK, A., ET AL. (2000): “Comparison of Health Inequalities between East and West Germany,” *European Journal of Public Health*, 10(4), 262–267.
- NOLTE, E., AND M. MCKEE (2004): “Changing Health Inequalities in East and West Germany since Unification,” *Social Science and Medicine*, 58, 119–136.
- NOLTE, E., V. SHKOLNIKOV, AND M. MCKEE (2000): “Changing Mortality Patterns in East and West Germany and Poland. II: Short-Term Trends during Transition and in the 1990s,” *Journal of Epidemiology and Community Health*, 54, 899–906.
- OECD (2005): *Pensions at a Glance. Public Policies Across OECD Countries*. OECD Publishing, Paris, France.
- PALME, M., AND S. SANDGREN (2004): “Parental Income, Lifetime Income and Mortality,” Mimeo, Department of Economics, Stockholm University.
- REHFELD, U. (2001): “Die Statistiken der Gesetzlichen Rentenversicherung. Zu Stand und Perspektiven des leistungsfähigen, vielgenutzten Berichtswesens.,” *Deutsche Rentenversicherung*, 3-4/2001, 160–186.
- REIL-HELD, A. (2000): “Einkommen und Sterblichkeit in Deutschland: Leben Reiche länger?,” Sonderforschungsbereich 504 Discussion Paper No. 00-14.
- SMITH, J. P. (2004): “Unraveling the SES-Health Connection,” IFS Working Paper 04/02.
- VDR (2004): *Statistik Rentenbestand am 31. Dezember 2003*. Verband Deutscher Rentenversicherungsträger (VDR), Frankfurt am Main, Germany.

**SONDERFORSCHUNGSBereich 504 WORKING PAPER SERIES**

Nr.	Author	Title
07-08	Daniel Schunk	The German SAVE survey: documentation and methodology
07-07	Hans-Martin von Gaudecker Carsten Weber	Mandatory unisex policies and annuity pricing: quasi-experimental evidence from Germany
07-06	Daniel Schunk	A Markov Chain Monte Carlo Multiple Imputation Procedure for Dealing with Item Nonresponse in the German SAVE Survey
07-05	Hans-Martin von Gaudecker Rembrandt Scholz	Lifetime Earnings and Life Expectancy
07-04	Christopher Koch Daniel Schunk	The Case for Limited Auditor Liability - The Effects of Liability Size on Risk Aversion and Ambiguity Aversion
07-03	Siegfried K. Berninghaus Werner Gueth M. Vittoria Levati Jianying Qiu	Satisficing in sales competition: experimental evidence
07-02	Jannis Bischof Michael Ebert	Inconsistent measurement and disclosure of non-contingent financial derivatives under IFRS: A behavioral perspective
07-01	Jörg Oechssler Carsten Schmidt Wendelin Schnedler	Asset Bubbles without Dividends - An Experiment
06-16	Siegfried K. Berninghaus Hans Haller	Pairwise Interaction on Random Graphs
06-15	Markus Glaser Philipp Schmitz	Privatanleger am Optionsscheinmarkt
06-14	Daniel Houser Daniel Schunk Joachim Winter	Trust Games Measure Trust
06-13	Markus Glaser Sebastian Müller	Der Diversification Discount in Deutschland: Existiert ein Bewertungsabschlag für diversifizierte Unternehmen?