Schmollers Jahrbuch 132 (2012), 323 – 357
Duncker & Humblot, Berlin

The Social Costs of Health-related Early Retirement in Germany: Evidence from the German Socio-Economic Panel

By Gisela Hostenkamp and Michael Stolpe

Abstract

Using data from the German Socio-economic Panel, we study how stratification in health and income contributes to the social cost of health-related early retirement, the balance of lost labour income and health benefits. On average, early retirees improve their health by almost two thirds of the loss suffered during the last four working years. We calibrate counterfactual scenarios and find keeping all workers in very good health, the highest of five categories of self-assessed health, would delay the average retirement age by more than three years and reduce the social costs by more than 20 percent.

JEL Classification: H55, I12, O15

1. Introduction

Summary. This study investigates how stratification in health and income contributes to the social cost of health-related early retirement, as evidenced in the German Socio-economic Panel (SOEP). We interpret early retirement as a mechanism to limit and partly reverse work-related declines in health, allowing poorer and less healthy workers to obtain a higher total discounted value of annuities from Germany's pay-as-you-go pension system. Our estimates indicate that within two to three years, early retirees recover on average almost two thirds of the loss of health status suffered during the last four working years before retirement, an effect we do not observe for workers retiring at the normal age. Balancing health benefits and labour income losses, we estimate net costs of almost € 60 billion in every year of the 1992 – 2005 sample period. Investments in new medical technology and better access to existing health services could help curb the need for early retirement and greatly improve efficiency. To value the potential gains, we calibrate counterfactual scenarios using an intertemporal model that is based on ex post predictions from stratified duration regressions for individual retirement timing. We conclude that eliminating the correlation between income and health decline would delay the average age of retirement by approximately half a year. Keeping all workers in very good health, the highest of five categories of self-assessed health, would yield a further delay of up to three years. Had this scenario been realized during our 1992–2005 sample period, we estimate the social costs of early retirement would have been more than 20 percent lower, not counting the direct social benefits from the counterfactual assumption of better health.

Motivation. As one of the first countries in Europe, Germany is making a serious effort to contain the costs of early retirement. It aims at raising the retirement age by making early retirement financially less attractive and by increasing the statutory retirement age from 65 to 67 years over the next two decades.[1] Today, only about 33% of workers retire at age 65 or after: in 2006 the average retirement age across all available retirement schemes was 60.9 years (Deutsche Rentenversicherung Bund, 2006). Previous attempts to quantify the costs of early retirement have focused almost exclusively on the expenditure side of the pension system and the implied losses in terms of GDP. In this vein, Herbertsson/Orszag (2003) predict a rapid increase in Germany's social costs of early retirement amid population aging – with output lost to early retirement reaching almost 13% of GDP in 2010 under the assumption of constant participation rates of elderly workers.

Our study focuses on bad health as a determinant of early retirement and quantifies the social cost of *health-related* early retirement from a social planner's perspective, consolidating aggregate income losses and health benefits that can be attributed to early retirement. We thus provide an estimate of the social value of investments that enable all workers to maintain very good health up to the statutory retirement age and beyond. Among various determinants of retirement, the literature has long recognized health as a key factor in the individual retirement decision[2] (see for example Dwyer/Mitchell, 1999; Larsen/

---

1 Countries with similar policies include Denmark where the normal retirement age will be raised from 65 to 67 in steps of six months per year beginning in 2024 and an automatic link to increases in life expectancy at age 60 will take effect after 2027. Ten other OECD countries have moved to defined-contribution pension plans without adjusting the normal retirement age. See Whitehouse (2007) for details.

2 An individual choice-based theory of retirement, required for meaningful welfare analyses, presupposes that the demand for retirement is determined ex ante – by factors exogenous at the time of decision-making, which may include the well-defined *objective* medical conditions that justify disability pensions under the law (Erwerbsminderungsrenten). Normally, the presence of such conditions will not automatically trigger early retirement, but merely offers the worker an opportunity to make an individual decision and proceed with the application and examination process. Ineligible workers may pursue other available pathways to early retirement, especially in the presence of health problems that are insufficiently severe for a disability pension or less well-defined than required by law. To identify bad health as a determinant of retirement timing, choice-based theories generally cannot rely (solely) on measures of eligibility for disability pensions that are determined on the basis of (changing) legal definitions ex post. To be testable, these theories must use explanatory variables that are observable in both retiring and non-retiring workers ex ante, such as *subjective* measures of self-assessed health ob-

Gupta, 2004). However, most studies have tended to neglect the work-related health hazards that early retirement *removes*, creating a welfare benefit that lowers net social costs. An exception is Vermeulen / Kalwij (2006) who estimate that labour force participation rates of elderly men would increase by about 12 % if everybody was in perfect health, a scenario which may help to establish an upper limit for the social costs of health-related early retirement.

These issues are difficult to address in structural retirement models, such as Gustman / Steinmeier's (1986) seminal model with a single representative agent, perfect consumption smoothing and limitless borrowing. As there is no difference between the social and private cost of health-related early retirement, these models do not provide a rationale for government policies. Moreover, without allowing substantial heterogeneity in preferences, they would not necessarily predict that work-related declines in health induce early retirement since workers' adjustment to relative work-leisure price changes would be subject to an income and a substitution effect in opposite directions. When declining health reduces the effective wage, workers *increase* their labour supply if the income effect dominates and *reduce* it if the substitution effect dominates. To account for the observed variation in retirement timing in a sample of US workers, Gustman / Steinmeier (1986) estimate parameters of a CES utility function that indicate wide variation in preferences for leisure, partly attributable to individual aging that lets the income-leisure indifference curves become steeper at an annual rate of 23 %.

Instead of preferences, our paper emphasizes the role of differential constraints and opportunities that the labour market assigns to workers when they are young. Workers' health is assumed to decline in a stratified pattern with strata defined on the joint distribution of health and income. This eliminates much of the theoretical ambiguity about the net effect on retirement timing because the relative size of substitution and income effects varies systematically across these strata. See Figure 1 for a stylized illustration in which the additional lifetime consumption that future wage income affords from a given age is plotted on the horizontal and planned time in retirement on the vertical axis. The reservation level of consumption a worker would have upon immediate retirement is normalized to zero; it depends on all sources of income in the absence of wage income, such as capital income, pensions and welfare payments. For workers in good health, starting with an initial allocation of leisure and working time in point A on budget constraint I, an unexpected increase in the relative price of consumption, as would be implied by an irreversible health shock, shifts the budget constraint to II and leads to an increase in labour supply and *less* time in retirement because the income effect (IE) dominates the substitution effect (SE). Workers with poorer health and lower productivity,

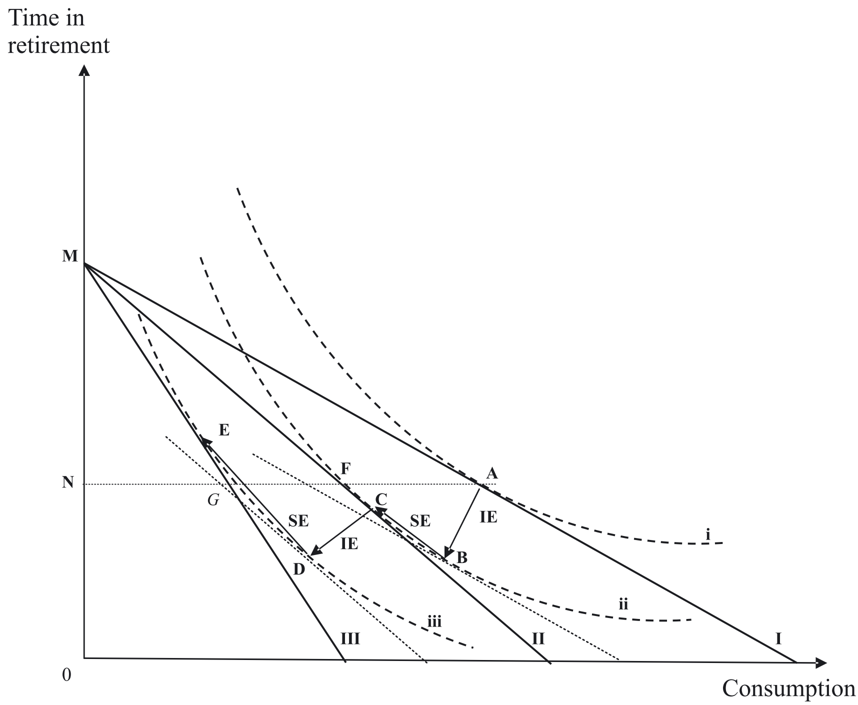
tained through surveys. Moreover, an exclusive focus on legal eligibility for disability pensions would bias the analysis by ignoring workers with impaired health that choose other pathways into retirement.

starting in point C, face a higher relative consumption price to begin with and the substitution effect of an unexpected further price increase dominates the income effect, resulting in *more* time in retirement, in point E on budget constraint III.

Either a loss in productivity or an increase in work-related mortality may trigger the unexpected consumption price rise shown in Figure 1.[3] A decline in general health is often accompanied by a rise in expected mortality, albeit in a non-linear fashion across strata as mortality increases faster when health declines from a low initial level of health. This exacerbates the relative work-leisure price change for a given health shock. Inter-temporal optimization within a lifecycle model implies that individuals expecting to be long-lived anticipate the need for greater wealth to finance more years in retirement and therefore want to retire later than those expecting to die early (Hostenkamp / Stolpe, 2008). Using nonparametric estimates of mortality rates by socio-economic status among male German pensioners, von Gaudecker / Scholz (2007) find that period life expectancy at age 65 rises almost linearly over a sizable part of the lifetime earnings distribution and varies by almost 50% or six years between the highest and the lowest of eleven socioeconomic groups in that study. Based on all public pensions and deaths in Germany during 2003, Shkolnikov et al. (2007) find that mortality varies by 60% between pension income quintiles, with period life expectancy ranging from 14.9 to 18.5 years.

Since the substitution effect in response to a given work-related health shock is larger among workers of low initial health, and health and income are correlated in the cross section of workers at any given age, we conjecture that the *average* substitution effect in the population of all workers rises with the variance in health and income, holding per capita income fixed. The variance rises within each cohort as it ages until about 60. However, the early-retirement option limits the extent to which workers allow their health to deteriorate and thus constrains the rate of mortality that is attributable to work-related health

3 The graph normalizes the maximum time workers expect to be able to choose in retirement to M, regardless of age, health or income. Without this normalization, poorer workers would face fewer years of farther life expectancy at any given age, but this would only change the scale of the graph in Figure 1. We acknowledge that poor health might be associated with a steeper slope of the relevant utility indifference curves that would indicate a lower marginal utility of consumption, as evidenced in Finkelstein et al. (2008). For US workers, Gustman / Steinmeier (1986) estimate that the onset of long-term health problems causes workers' instantaneous income-leisure indifference curves to increase in slope by 132%. This would reinforce the impact of declining labour productivity on the substitution effect as less healthy workers would choose a lower level of lifecycle savings, even if life expectancy were held fixed. Finally, we might extend the graph to show that a given health shock need not only change relative prices, but may also shift the budget constraint towards the origin, reflecting lower life expectancy. Again, while this shift would increase the income effect, it would not overturn our conclusion that the substitution effect is more likely to dominate among poorer and less healthy workers.



Note: This graph is meant to represent a stylized illustration of the income and substitution effect that workers face at a given point in time – after the relative price of consumption financed by labour income has unexpectedly increased, as would be implied by an irreversible health shock. The graph does *not* represent a directly testable model. For workers in good health (starting point A) the income effect (IE) dominates the substitution effect SE: in response to a higher relative price of consumption goods in terms of forgone leisure, these workers increase labour supply and consume lower levels of leisure. For workers in poorer health (starting point C) the substitution effect dominates the income effect: they choose higher levels of leisure by retiring early. For ease of exposition, the reservation level of consumption that the worker would have in case of immediate retirement is normalized to zero; it depends on all sources of income the individual may have in the absence of wage income, such as capital income, pensions and welfare payments.

Figure 1: Income- and Substitution Effects in Retirement Planning for Different Initial Levels of Labour-income-financed Consumption and Health

hazards. For a given degree of stratification, inequality in life expectancy will be lower than it would be without the option to retire on health grounds, which we therefore interpret as a social insurance mechanism that limits the losses of those whose health deteriorates fastest. Although the effect on people's individual incentives to invest in their human capital ex ante is theoretically ambiguous, as we will show in section 2, there is likely to be a classic equity-efficiency trade-off. Without early retirement, the economy's total output could be larger, health spending could be higher and society could more easily bear the often large fixed costs of developing and introducing new medical technologies.

To help quantify the potential welfare gains from investments that reduce or eliminate the work-related deterioration of health, we compare the social cost of early retirement in the presence of stratification in health and income with the hypothetical social cost in its absence. We define the absence by imputing counterfactual health states, such as very good health, the highest of five self-assessed health categories recorded in the German Socio-economic Panel (SOEP). To implement this approach, we investigate the timing of early retirement in the context of the age-group-specific health gradient, defined as the positive contemporaneous correlation between workers' income and health within each cohort over non-overlapping five-year age intervals. The health gradient is an empirical regularity that has been observed in many countries and across the whole range of income classes. Much recent research points to underlying third factors, such as the long reach of events in childhood, parental income and education and the intergenerational transmission of health and wealth as the ultimate explanation (Case et al., 2002; Currie / Stabile, 2003; Cutler et al., 2008). Without implying contemporaneous causality, recent empirical studies, such as Deaton / Paxson (1998) for the US and Hostenkamp / Stolpe (2006) for Germany, show that the health gradient tends to be steeper in older age groups until about age 60. Our own empirical findings suggest that workers' individual positions relative to the rising cohort-specific health gradient are persistent as they age. At each point in time, workers' propensity to retire early is hence negatively correlated with their wage income and stratification accelerates the decline of the opportunity costs of early retirement with age for those with poor health and poor earnings prospects in the labour market.[4]

By raising the stakes and revealing a coordination problem, persistence and stratification provide a rationale for greater public investments in medical technology, health services research and healthcare to counter work-related shocks to health. Those locked into a position of relatively poor health and low income face a specific borrowing constraint[5] as they cannot hope to be financially rewarded by higher wages for individual investments in their health, although society would benefit. By contrast, high mobility of workers, such that they often change their relative position as they age, would strengthen the case for setting individual incentives, including further changes in pension rules to increase flexibility and improve the private incentives for early retirees to return to work when their health recovers. Germany's disability pension reform of 2001 can actually be interpreted as a significant step in this direction: The first disability pension may now only be granted for a limited time and is subject to

4 In a more recent study, based on the 2008 cross section of new pension claims in Germany's statutory system, Hagen et al. (2011) confirm that workers with poor educational attainments are far more likely to become a disability pensioner.

5 Using data from the US Panel Study of Income Dynamics, French (2005) demonstrates the empirical significance of borrowing constraints for retirement behaviour amid uncertain future health and wages.

review by government-employed doctors after a pre-specified number of years, often annually.

The remainder of this paper is organized as follows. In section 2, we start with a definition of the social cost of early retirement and describe alternative specifications of the counterfactual absence of the health gradient. Section 3 describes our data and presents duration regressions for the timing of early retirement by individual workers in Germany. Section 4 explains how we use these regression results to make counterfactual predictions of retirement timing, presuming workers' relative position on the health gradient is persistent. Section 5 finally quantifies the social costs of health-related early retirement in monetary terms. Section 6 concludes.

2. The Problem of Social Costs in Health-related Early Retirement

The components of social costs. We distinguish between gross and net social costs and define the latter as the difference between the gross social and the private costs of health-related early retirement. To obtain measurable aggregates, we define the gross social costs as the retiring worker's full output plus the pension bill to society. The private costs are limited to lost wage income minus any private benefits, such as pension payments, leisure time and gains in health and longevity. The gross social costs exceed the private costs when the retiree is still in relatively good health, but they eventually fall below the private costs when the retiree's health declines: While healthy workers' average output generally exceeds their wage income, assuming wages equal workers' marginal product, the private costs do not account for the possibility that declining health lets a worker's productivity drop below the wage level.[6] Our empirical focus is on the *net* social costs, where we assume for simplicity that the discounted sum of pension payments equals the present-valued pension bill to society and the discounted sum of the private value of early retirees' gain in leisure equals the discounted value of their lost contribution to GDP. This seems a reasonable approximation since labour market equilibrium implies wages equal the marginal value of leisure; and healthier retirees, from the higher strata of workers, can be assumed to value leisure more than do those in poor health.[7]

6 We note that the private costs tend to be reduced by mortality-contingent claims, such as pension annuities and the benefits of health insurance with guaranteed renewal, whereas the social costs may be exacerbated by the excess burden from a declining labour supply in response to rising pension and health system contributions when people increase their time in retirement. We think these second-order effects are small enough to be ignored here.

7 Finkelstein et al. (2008) provide evidence that changes in health affect the marginal utility of consumption. Since the consumption of many goods and services requires lei-

Alternatively, we may interpret the balance between health benefits and lost wage income as an empirical assessment of the net payoff from the implicit insurance that the option of early retirement provides against the continuation of work-related health declines. As this insurance protects a minimum level of individual health and a maximum level of social health inequality, we can think of it as similar to Shiller's (2003) concepts of livelihood and inequality insurance for income. For an illustration, consider once more Figure 1 where the absence of an early retirement option would constrain all workers to spend N units of time in retirement, as indicated by the horizontal line from N to A. Those in the poorest health state would be at point G, below the utility indifference curve iii that goes through point E, which can be reached by early retirement. Moreover, because they would work longer these workers would suffer declines in health, represented by a rotation of budget constraint III on M towards the origin, which would exacerbate the welfare loss. At the same time, a strict mandatory retirement age would prevent those in good health from increasing their labour supply to point C. Instead, they would be constrained at point F, below the utility indifference curve that goes through C.

Figure 1 thus indicates that the early retirement option does not only provide some protection against further work-related declines in health, but also implies some pooling of resources and sharing of the financial burden, especially when those in good health are allowed to extend their working lives. Suppose that only the change from budget constraint II to budget constraint III is due to work-related health deterioration and that all workers participate in a pay-as-you-go pension system. In this case, a rotation of budget constraint I towards II can arise simply because pension payments to early retirees increase the payroll tax for active workers, lowering the relative price of leisure in terms of consumption goods and creating incentives for the healthy to work more. The implied insurance might therefore be efficient if work-related risks to health were truly independent across individuals. The insurance aspect could then even improve the incentives of young workers to save and invest in their human capital as they would anticipate the option to constrain the impact of work-related risks to health and longevity should they materialize at some higher age. Expecting better health and a longer life, they will aim at a higher level of lifecycle income and wealth (Becker, 2007; Finkelstein et al., 2008; Hostenkamp / Stolpe, 2008).

However, stratification on the joint distribution of health and income implies that work-related risks to health are *not* independent. Instead, there is a zero-sum aspect to workers' individual human capital investments so that, from an ex ante point of view, the size of long-term work-related risks to health is negatively correlated with those of workers entering strata at the opposite end of the distribution. This prediction follows from the human capital theory of health

sure as a complement, we conjecture that its marginal utility will also depend positively on health.

demand (Grossmann, 1972; Muurinen / Le Grand, 1985) in Case and Deaton's (2005) version which implies systematic variation in workers' rates of health decline when medical technology is imperfect and healthcare cannot fully repair all work-related shocks to health. The rate of health decline depends primarily on the type of jobs, distinguished by their demands on health, i.e., the size and frequency of irreparable health shocks. Workers make human capital investments to compete for the best jobs in terms of income and health, a process in which the superior human capital of some may relegate others to less favourable jobs, potentially making part of their human capital investments obsolete. Individual investments can thus create negative pecuniary externalities. The labour market is characterized by assortative matching (Becker, 1973, 1974, 1993) that places young workers in different strata of health development and in effect denies many workers their full potential over the lifecycle – especially workers "selling their bodies" (Case / Deaton, 2005, 194) in hard manual jobs or "selling their mental health" in jobs with little span of personal control over the content and organization of one's work. Becker (2007) and Hostenkamp / Stolpe (2008) discuss a number of complementarities that stabilize stratification over the lifecycle, in particular the complementarity between workers' health and education and between health investments at different ages. Initial differences in health are reinforced and differential rates of health deterioration across strata make individual positions relative to the cohort-specific health gradient persistent as its slope rises with age.

Although the insurance implied by the early retirement option may still help to improve the ex ante incentives, it is no longer *dynamically* efficient. It neither mobilizes nor rewards the necessary investments in health and education to overcome the problem of stratification in income and health. On the contrary, as widespread early retirement depresses GDP, society will find it more difficult to cover the fixed costs of health investments, such as new medical technology, that are subject to economies of scale. Those in good health might be willing to subsidize health investments for those in poor health if these could make a credible commitment not to retire early and instead to continue contributing to the funding.[8] But without coordination and enforcement, an efficient deal will not be made (Hostenkamp / Stolpe, 2008).

Empirical implementation. To assess the *net* social costs empirically, we subtract from the gross social costs workers' private costs, including the value of

8 The role of health investments targeted at poorer workers in achieving equity and efficiency in retirement incentives is highlighted by recent empirical studies that show these workers die so much earlier that their expected total lifecycle pensions are less than actuarially fair even if pension payments are actuarially fair for the average population (Frijters et al., 2005; von Gaudecker / Scholz, 2007; Jürges, 2005). In a similar vein, Siddiqui (1997) estimates that the bad health of many workers reduces the effectiveness of pension adjustments that are actuarially fair for the average worker in raising the average German retirement age by one fourth, namely from 2.4 to 1.8 years.

gains in leisure and an estimate of workers' willingness-to-pay for the mortality reduction that can be attributed to health benefits observed only after early retirement. Following Becker (2007, 385), four fifth or more of this willingness-to-pay should be attributable to the benefit of enjoying leisure instead of work as the vast majority of the statistical value of life does not stem from earnings, but from leisure time, and differences between average and marginal utilities.[9] We then compare the observed labour income losses and health benefits from early retirement with a counterfactual scenario in which the health gradient is absent. In fact, we specify two variants of the counterfactual scenario. One is based on the assumption that all workers enjoy the same evolution of the health status as is observed for the average worker in the highest income quintile. The other is based on the assumption of very good health for all.

To this end, we need to estimate a dynamic model that allows for state dependence in the accumulation of health shocks over time, with states identified by workers' positions relative to the health gradient, not by health or income alone. Early retirement would not provide insurance of individual human capital investments if the evolution of workers' health were unrelated to their lagged positions relative to the age-group-specific health gradient.[10]

---

9 See Becker (2007, 385) for a back-of-the-envelope calculation of the value of the statistical life for a healthy American. Assuming an average annual income of $40,000 and 1.8 times as much time spent in unpaid household production as in formal work, after subtracting 68 hours for sleep and maintenance per week, Becker values full income at $110,000 per year. He adjusts this estimate upward to account for the concavity in the single period utility function which implies average utility of additional years of life exceed the marginal utility of consumption, and gets $220,000 of adjusted full income. Discounting this at 5% yields a value of life at about $4.4 million. This is more than four times the present value of the lost earnings from early death, which are merely $1 million. We ignore the possibility that changes in health affect the marginal utility of consumption (Finkelstein et al., 2008) or the utility of leisure.

10 Our approach therefore differs from Halliday's (2008) analysis of the evolution of health over the lifecycle. He uses data on self-reported health status from the US Panel Study of Income Dynamics to estimate first-order Markov transition probabilities without including a strata-identifying variable that would help to explain the persistence of ranks assigned by assortative matching. Halliday's (2008) statistical model allows for two sources of persistence, namely unobserved heterogeneity, interpreted as workers' intrinsic ability to cope with health shocks, and state dependence, the degree to which the ability to cope depends on health status. We argue instead that workers' position relative to the health gradient determines their ability to cope. As his results indicate a large degree of individual heterogeneity, Halliday (2008) argues that much of what determines health in adulthood can be traced back to childhood (Case et al., 2002; Currie/Stabile, 2003). Using data from the US Health and Retirement Study, Heiss et al. (2008) extend this line of research and include information on socio-economic attributes, such as educational attainment, to condition the estimates of health transition probabilities and simulate some of the implied health dynamics. Confirming expectations in line with our theory of the health gradient, they find people with low educational attainments, defined as eight years of schooling, are several times more likely to be in poor health at age 50 and beyond than people with high educational attainments, defined as 16 years of schooling.

With few exceptions, the literature has neglected the possibility that early retirement may slow down or even halt work-related declines of health and that workers value these benefits. Börsch-Supan / Jürges (2006) are among the first trying to measure the intangible benefits from early retirement in terms of wellbeing, but do not assign monetary values to these gains. However, several studies (Jürges, 2005; Frijters et al., 2005; Rehfeld, 2006) show self-rated health to be a good predictor of subsequent mortality, and we can use these estimates to quantify the impact of health improvements on mortality risk. In addition, there is a vast literature on the empirical willingness-to-pay (WTP) for mortality risk reductions, known as the statistical value of life, so that we can indirectly infer the value of health improvements after early retirement.

Since we assume GDP losses and gains in leisure to cancel out, we calculate the net social costs of early retirement for a given year as the discounted present value of foregone labour income minus the gains from mortality risk reductions. Aggregating across N workers, indexed by i, who retire at some individual age A_i, the social costs C can be written as

(1)

$$C = \sum\nolimits_i^N \left\{ \sum_{t=A_i}^{65} \frac{Y_i}{(1+\rho)^t} \times \Pr(s_{it}|A_i,\, G,\, R,\, H) - \sum_{t=A_i}^{100} \frac{WTP}{(1+\rho)^t} \times \Delta \Pr(s_{it}|A_i,\, G,\, R;\, H) \right\}$$

where Y_i is the annual gross labour income[11] at age t, assumed constant for all future years; ρ is a time preference rate equal to the 3% level that the World Bank and other international organizations apply for developed countries; and $\Pr(s_{it}|A_i, G, R, H)$ is the conditional probability to survive at least t more years, given survival until age A, gender G and region of residence R.[12] In addition, the survivor function depends on a worker's health status H: Better than average age-specific health is assumed to shift the survivor function upward, to match the mortality rate as if the person were of average health and younger, while worse than average health is assumed to shift the survivor function downward, as if the person were older.[13] Throughout, we assume the survivor function is independent of the survival of spouses. The second sum in the curly brackets is the discounted present value of willingness-to-pay for mortality risk

11 For lack of better data, we do not include the full value of workers' output if they did not retire as a cost of early retirement. Total output will generally exceed the marginal product of labour because most workers are infra-marginal, but gross wages are best assumed equal to workers' marginal product.

12 The conditional survival probabilities are computed from the standard life tables of the Statistisches Bundesamt (2006) which provides this information separately for both genders and the regions East and West Germany.

13 In a similar vein, Gustman / Steinmeier (1986) estimate that the onset of long-term health problems is equivalent to an increase of almost four years in workers' age.

reductions, where WTP is obtained from the literature on the value of a statistical life in Germany, $\Delta \Pr(s_{it}|A_i, G, R, H)$ is the change in survival probability that can be attributed to health-related early retirement, assuming the reduction of mortality risk will be effective beyond the normal retirement age until a hypothetical maximum lifetime of 100 years – the last year for which age-specific mortality rates are available.

Constructing counterfactual scenarios. Defining the most appropriate counterfactual scenario to determine when workers would retire in the absence of the health gradient is not trivial because equity in access to health need not imply equality in health. A social justice approach (Bommier/Stecklov, 2002) suggests that the ideal distribution of health would be reached if access to health were not determined by income or socioeconomic status. Thus, knowing a person's health status does not help to predict income, and vice versa. However, several different scenarios with distinct normative implications would be consistent with this condition. *First*, there would be no health gradient if there was no income inequality. In this vein, Deaton (2001) suggests an argument for income redistribution based on the observation that a concave relationship between individual health and income will cause greater income inequality to be associated with lower average population health: Since health is not transferable, inequality aversion calls for mean-preserving transfers of income.

Second, there would be no health gradient if the distribution of health states across individuals were independent of income. Kakwani et al. (1997), concerned with aggregate measures of income-related health inequality, define "avoidable" inequality as the difference between the observed health concentration index and a standardized concentration index in which the distribution of health is age- and gender-standardized to account for these "unavoidable" demographic factors. However, it is not clear why the age- and gender-specific concentration of ill health should be unavoidable (Chotikapanich et al., 2003). Van Doorslaer/Koolman (2004) instead use the 1996 cross section of the European Community Household Panel to define "excess" inequality by relating the health concentration index in one country to the degree of concentration in the country with the lowest inequality. They argue that health inequalities should primarily be addressed by health policy, not income redistribution, as countries with the low income inequality are not always the countries with the lowest income-related health inequality.

Third, there would be no health gradient if everybody had the same health status, irrespective of whether that status was excellent, fair or bad. Vermeulen/Kalwij (2006), for example, assess the importance of being healthy for individual retirement behaviour by quantifying a counterfactual scenario where everybody is in perfect health and labour market participation rates are more than 12%age points higher among German men, and about 8%age points among women. We think equal health for all is the least ambiguous scenario to quantify. By contrast, constructing counterfactual scenarios on the basis of no income inequality or on the basis of statistical independence of health and income is more difficult as these assumptions are specific only at the population level, but do not specify how counterfactual health states should be assigned to individual workers.

3. Estimating Individual Retirement Behaviour in Germany

Construction of the health gradient. Because measuring health is difficult and structural lifecycle models with endogenous health cannot be estimated in a reliable way, we follow Adams et al. (2003) and other recent authors who have estimated relatively simple reduced-form models. Our study is based on an *unbalanced* panel drawn from 13 consecutive annual waves of the SOEP starting in 1992 when our preferred health measure was first collected. The testable prediction of a systematic correlation between the individual position relative to the health gradient and retirement timing provides the empirical basis for our approach to identifying the social costs of health-related early retirement. The health gradient is constructed separately for workers within each five-year age interval using an OLS regression of a cardinalized measure of health on the natural logarithm of gross income. This enables us to assign a cardinal measure of individual positions relative to the age-group-specific health gradient to each worker by taking the orthogonal projection of his or her individual position relative to the OLS regression curve onto the income axis in the age-group-specific health-income diagram. We then subtract the age-group-specific median value to eliminate general aging-related changes in health or income. In a last step, we construct a discrete variable of workers' relative position on the health gradient defined by quintiles. The variance of individual positions increases as cohorts age and, with declines in health related to the type of work and stratified by income class, workers' quintile positions are persistent over time.

We use the variable *self-rated health status*, a subjective measure of health based on answers to the WHO-recommended question "How would you describe your current health?" on a five point ordinal rating scale from *very good* coded as 1 to *very bad* coded as 5. As respondents' self assessment of overall health is likely to include physical as well as mental aspects, we prefer this measure over more objective, yet often incomplete measures of health. Before constructing the health gradient, we use the empirical normal transformation, assuming a latent variable of health with a standard normal distribution (Jürges, 2005; 18–19), to obtain a cardinalized health measure.[14]

[14] This transformation assigns the values 1.577 for *very good*, 0.431 for *good*, –0.540 for *fair*, –1.373 for *bad* and –2.196 for *very bad* self-rated health categories.

The second variable needed to construct the health gradient is income. The SOEP provides continuous measures, recorded in euro, of nominal individual labour income and of total household income.[15] We transform these variables using the OECD consumer price deflator so that real income is expressed in prices of 2001. For regression analyses, we use the natural logarithm of the variable *post-government* income, which is defined as the combined total household income of all household members after taxes and transfers. In addition, we run regressions in which household income is replaced by equivalent income defined as:

$$\textit{equivalent income} = \frac{\textit{household income}}{0.5 + 0.5 \times (\textit{household members} - \textit{children}) + 0.3 \times \textit{children}}\,.$$

To identify the exact beginning of workers' retirement, we choose the month when a person first declares him- or herself retired as this is the most exact variable available in the SOEP. This variable, moreover, may be less affected by the justification bias that may arise if respondents understate their health in order to justify early retirement in a socially acceptable way. In the SOEP, respondents are not asked to give a reason for retirement, such as award of a disability pension (Erwerbsminderungsrente), and health and retirement variables are not recorded in the same questionnaire, so that conscious misreporting seems unlikely.

Our regressions also include variables that were significant explanatories of retirement timing in previous studies: the natural logarithm of imputed rent for owner occupied housing,[16] to control for wealth effects; the number of years in education, to measure a person's attainment beyond the legally required minimum; a dummy variable indicating whether a person is a public sector employee, to account for differences in the pension systems for private and public employees; and dummies for gender, region of residence, marital status and unemployment, which we all obtain from the CNEF equivalent file of the SOEP.

As the decision to retire is essentially a dynamic inter-temporal decision problem, the natural empirical approach is duration analysis. However, duration analysis cannot account for longitudinal weighting that requires person-specific time-varying weighting factors. To account at least for differences in initial sampling probabilities across subgroups, we use the cross sectional weights of the first year that an individual is observed to weight the subsequent observations. This might lead to a slight over-weighting of individuals that participated from the start in 1992, relative to individuals of subsamples included later, since we cannot examine differences in panel attrition rates across subgroups. Selection bias, due to the lower attrition of workers in good health, may render the estimated hazard to retire on health grounds too small.

15 In the so-called CNEF equivalent file.

16 For exact imputation procedures see Frick / Grabka (2000).

Table 1

Descriptive Statistics (unweighted)

Variable	Definition	Mean	Std. Dev.	Min.	Max
Self-rated health status (SRHS)	Ordinal health variable with 1 = very good to 5 = very bad health	2.754847	.9168782	1	5
Self-assessed health (SAH)	Continuous health variable after transforming SRHS using the empirical normal transformation	–.2594927	.855776	–2.195888	1.577415
Household income	Natural logarithm of household total income after government transfers indexed for inflation (including labour and asset income, private and public transfers)	10.29106	.6339893	0	14.14139
Household equivalent income	Natural logarithm of HH_income adjusted for equiv. household size	9.744243	.5614186	0	13.73592
Housing wealth	Natural logarithm of owner occupied housing wealth imputed by DIW	1.503461	1.810956	–6.214608	6.361366
Health gradient	Cardinal position relative to the median on the age-group-specific health gradient	–5.88e-06	.6013779	–10.34613	4.108216
Quintile position	Quintile position on the health gradient with 1 = lowest position; 5 = highest position (ordinal)	3.049581	1.394412	1	5
Employment status	Employment Status dummy with 0 = working ; 1 = non-working	.3469474	.4760012	0	1
Years of education	Years of education exceeding a 7 year minimum	4.981746	2.667117	0	11
Public sector	Dummy for public sector employment 0 = private sector 1 = public sector	.1602433	.3668325	0	1
Eastern residence	Dummy for East German Residence 0 = West 1 = East	.2562629	.4365709	0	1
Gender	0 = male 1= female	.5068003	.4999563	0	1
Married	Dummy for marital status 1 = married 0 = other	.8011029	.3991724	0	1
Year of birth		1945.681	9.622879	1912	1964

Total No. of Obs. = 186947; No. of Obs. end before age 40 = 86618; No. of Obs. beginning after retirement entry = 25725; No. of Obs. remain. representing = 74604 (total), 5.22437 (mean), 1(min), 5 (median), 13 (max); No. of Subjects = 14280 (total) and 426 (mean); Subj. with gap = 4; Time on gap if gap = 1704 (total), 426 (mean), 394 (min), 411 (median), 488 (max); Time at risk in days = 26122296 (total)), 1829.292 (mean), 14 (min), 1612 (median), 4962 (max). The consumer price index used is from the OECD: 1992 = 83.7; 1993 = 87; 1994 = 90.1; 1995 =92.6; 1996 = 94.1; 1997 = 95.3; 1998 = 97.1; 1999 = 97.9; 2000 = 98.6; 2001 = 100; 2002 = 102 ; 2003 = 103.4; 2004 = 104.5; 2005 = 106.2.

To account for unobserved heterogeneity, we estimate stratified duration models that allow the baseline hazard to differ between strata, defined on the basis of workers' quintile positions relative to the age-group-specific health gradient. In the context of our parametric duration analyses, this means that the shape parameter p is allowed to differ between strata but the distribution is confined to be of the Weibull family. This allows for groupwise time-varying heterogeneity due to unobserved persistent health shocks between, but not within strata, so that we can explore some basic implications of the human capital model of health demand, identify differences in retirement timing between workers in different strata and make out-of-sample predictions of the mean time to retirement.

Results of duration regressions. Table 2 reports our estimates of the Weibull model, stratified on workers' quintile position relative to the age-group-specific health gradient, using the lowest quintile as the base category. Positive coefficients indicate that higher covariate values are associated with a greater hazard to retire, and vice versa. A prior comparison of the log likelihoods from stratified and non-stratified regressions, available from the authors on request, showed that stratification significantly improves the explanatory power of the model. All dummy variables for quintile position turned out highly significant with the expected signs. The scale and the shape of the hazard function are significantly different from the base category in each of the other strata. The regressions also show that most of the variation can be explained by workers' quintile position, and that variation within strata is satisfactorily accounted for by the other significant covariates, as our continuous measure of workers' position relative to the health gradient becomes insignificant once quintile position is controlled for.

Our stratified Weibull regressions confirm that the hazard of early retirement is highest in the lowest quintile position relative to the health gradient. The estimated effect of moving from the lowest to the second lowest quintile is particularly large. In models using unadjusted household income, the hazard to retire is smallest in the highest quintile, whereas in models using income per equivalent household member, the hazard is smallest in the second highest quintile. Based on the Akaike information criterion, which is proportional to the absolute value of the likelihood function plus the number of estimated parameters, models using income per equivalent household member perform best.

The estimated coefficients for the other covariates are also robust across different model specifications and have the expected signs. More specifically, we find that public-sector employment, unemployment for males and eastern residence for females significantly increase the hazard to retire, consistent with the view that financial variables are important determinants of retirement behaviour at the individual level. However, continuous measures of household income are always insignificant in our stratified regressions, so that their impact seems to be fully accounted for by the strata identifying variable. Summing up, we find that health, education and financial variables have a significant effect on the retirement behaviour of both men and women.

Table 2

Results of Stratified Parametric Duration Models for Early Retirement

Coefficients	Household income	Per equivalent HH member
Self-rated health	**–0.24**	**–0.23**
$P > \|z\|$	0.000	**0.000**
Household income	**–0.04**	**–0.01**
$P > \|z\|$	0.542	0.938
2nd quintile	**–7.44**	**–10.09**
$P > \|z\|$	0.004	0.000
3rd quintil	**–15.39**	**–15.90**
$P > \|z\|$	0.000	0.000
4th quintile	**–12.83**	**–18.09**
$P > \|z\|$	0.000	0.000
5th quintile	**–17.05**	**–17.18**
$P > \|z\|$	0.000	0.000
Cons	**81.81**	**79.12**
$P > \|z\|$	0.000	0.000
2nd quintile *ln_p*	**0.31**	**0.41**
$P > \|z\|$	0.003	0.000
3rd quintile	**0.56**	**0.58**
$P > \|z\|$	0.000	0.000
4th quintile	**0.48**	**0.64**
$P > \|z\|$	0.000	0.000
5th quintile	**0.60**	**0.61**
$P > \|z\|$	0.000	0.000
Cons	**0.82**	**0.79**
$P > \|z\|$	0.000	0.000
No. of subjects	14280	14280
No. of failures	2867	2867
Time at risk	26122296	26122296
Log likelihood	–2973046.9	– 2880578.9
No. of obs	74604	74604
Wald chi^2	454.47	440.83
Prob > chi^2	0.000	0.000
Akaike Information Criterion	5946127.8	5761191.8

Note: The coefficients of our additional control variables – dummies for housing wealth, male, no employment of males and females, respectively, public sector employment, sex, married females, and Eastern residence as well as the year of birth and years of education – are available from the authors on request. Standard errors adjusted for clustering on individual persons.

4. Predicting Individual Retirement Behaviour

Based on ex post predictions of hazard rates and mean time to retirement, this section examines the persistence of workers' positions relative to the age-group-specific health gradient, using Markov chains, and quantifies counterfactual scenarios in which the health gradient is absent. Based on our stratified regression estimates with household income and health as separate covariates, the hazard function is

$$(2) \qquad \hat{h}\left(t_j \middle| x_j\right) = \exp(\beta_0)\exp(\ln(p)) \cdot t_j^{\exp(\ln(p))-1}$$
$$\exp\begin{pmatrix} +\beta_1 srhs + \beta_2 public + \beta_3 unempl + \beta_4 edu \\ +\beta_5 sex + \beta_6 married + \beta_7 east + \beta_8 birthyear \end{pmatrix}$$

where β_o and $\ln(p)$ vary by quintile position. Household income enters our analysis via the definition of quintile positions relative to the age-group-specific health gradient. Assuming constant covariate values, the mean time to retirement is the expected value of survival time, or the first moment of the distribution of survival time, conditional on covariate values $x_j : \mu_{Tj} = E\left(T_j \middle| x_j\right) = \int_0^\infty f\left(t \middle| x_j\right) dt = \int_0^\infty S\left(t \middle| x_j\right) dt$. We use these predictions to calibrate the welfare effects of health-related early retirement without attempting to account for time-varying covariates at the individual level.[17] The implied prediction error will be greater, the more the main covariates of interest in our analysis, workers' health and their individual position on the health gradient, change in reality.

Assuming a health status or quintile position that is *worse* than the true value would result in an under-prediction of the mean time to retirement, because both better health and a better individual position on the health gradient lower the predicted hazard to retire in each point in time, thus increasing predicted time at work. By contrast, assuming a health status or quintile position that is *better* than the true value would result in an over-prediction of the mean time to retirement. To obtain a realistic corridor for the estimated time to retirement of each individual, we perform two sets of ex-post predictions. In the first, we set all covariates constant at the values observed at age 40, when workers' time at risk to retire is set to begin. In the second, we set the covariate values constant as of the moment of retirement; in virtually all cases, workers' health is

[17] It is beyond the scope of this study to estimate how covariate values change with time at the population level although in principle these changes could be incorporated in our predictions. Yet, variations in the rate of covariate changes at the individual level could still not be accounted for. Our predictions thus may not be able to fully account for behavioural changes that may be induced by large exogenous changes in the institutional environment, such as new laws and regulations that the government may introduce in the future.

better at age 40 than at the moment they retire. Table 3 shows the mean time to retirement across all observations for these two scenarios. The difference is about 155 days. Expected retirement age is 59.02 years when covariates are set constant at age 40. It is 58.60 when covariates are assumed constant as of the moment of retirement for each worker. Our ex post predictions are thus very close to the actual average retirement age for the entire sample, which is 58.9 years.[18]

As workers' quintile position is defined relative to the age-group-specific health gradient, whose slope increases with age, there is bound to be persistence over time. And the more persistent workers' positions, the more reliable can be our long-term predictions of retirement timing. Put differently: Since workers' quintile position is estimated to be the dominant influence on the retirement hazard, as shown in the bottom panel of Table 2, our prediction error will increase with workers' propensity to move from one quintile to another. To obtain an empirical measure of persistence, we estimate a discrete-time Markov chain and compare the size of entries on the main diagonal with off-diagonal entries. The states of the Markov chain are defined by quintile positions so that our estimates give information on workers' intra-distributional mobility between quintiles on the health gradient. We use fractile Markov chains to ensure every discrete state contains the same number of subjects. Nonetheless, the distribution of workers in our sample is not exactly equal across quintiles because the different sampling probabilities of each individual are taken into account using the person expansion factors provided by the SOEP.

Results of Markov chain estimation. Table 3 presents estimates of first-order time-stationary transition probabilities over periods of one and five years for the population 40 plus.[19] The first panel gives the one-step annual transition matrix, whose (i, j) entry is the conditional probability that an individual has transited from quintile i to quintile j after one year. Persistence is greater in quintile one and five, exceeding 68 % and 76 % respectively. Yet considerable intra-distributional mobility that is evident in the three middle quintiles, where the entries on the main diagonal fall below 50 %, leads to a long run stationary distribution, reported in the first panel along with the sample distribution, that does not show any concentration of probability mass at the extremes.

18 Note that the Deutsche Rentenversicherung (2007) reports an official average retirement age in Germany of 60.9 years in 2006, reflecting the fact that the average retirement age increased slightly during the observation period.

19 Ignoring any information about transition probabilities which may be contained in the initial probability distribution and assuming the transition probabilities to be invariant with respect to time as well as gender, we use the Maximum Likelihood estimator $p_{ij} = h_{ij}/h_i$ where h_{ij} denote the observed frequency of transitions from state i to state j and $h_i = \sum_j h_{ij}$, to determine the transition probabilities. The transition matrix can then be used to calculate a long run stationary distribution using the Chapman-Kolmogorov equation as described in Osaki (1992).

Table 3

Five State Fractile Markov Estimates for Transition Probabilities between Quintiles on the Health Gradient for Population 40+

Observations	Transition end state Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
	First order, time stationary estimates of the one-year transition probabilities				
16306	**0.682**	0.229	0.06	0.021	0.008
19440	0.189	**0.494**	0.24	0.064	0.012
18906	0.052	0.234	**0.439**	0.232	0.043
18729	0.021	0.065	0.221	**0.503**	0.19
19349	0.008	0.014	0.039	0.171	**0.768**
Stationary Distribution	0.174	0.204	0.201	0.202	0.219
Sample Distr.	0.176	0.21	0.204	0.202	0.209
	First order, time stationary estimates of the five-year transition probabilities				
7354	**0.525**	0.296	0.122	0.042	0.015
9285	0.204	**0.371**	0.277	0.122	0.027
8506	0.103	0.231	**0.319**	0.266	0.081
7886	0.049	0.106	0.224	**0.378**	0.243
6393	0.024	0.037	0.077	0.216	**0.645**
Stationary Distribution	0.164	0.202	0.206	0.213	0.215
Sample Distr.	0.187	0.236	0.216	0.2	0.162
	One-year transitions iterated 5 times				
	0.303	0.273	0.201	0.139	0.084
	0.229	**0.252**	0.219	0.177	0.123
	0.17	0.219	**0.22**	0.208	0.182
	0.12	0.175	0.204	**0.233**	0.268
	0.072	0.117	0.163	0.243	**0.405**

Age-group	Boundary 1	Range 2	Range 3	Range 4
43	–0.385	0.272	0.230	0.294
48	–0.419	0.293	0.254	0.317
53	-0.423	0.303	0.249	0.314
58	**–0.430**	0.304	0.250	0.318
63	–0.408	0.282	0.250	0.335
68	–0.359	0.246	0.222	0.305
73	–0.332	0.217	0.196	0.304
78	–0.337	0.233	0.218	0.355

The second panel gives estimates of first-order five-year transition matrices, where the entries on the main diagonal are lower, indicating much lower persistence over longer periods of time. Nonetheless, the one-year transition estimates are qualitatively confirmed. To illustrate consistency of our short- and long-run estimates, we iterate the one-year transition matrix five times and report this in the third panel. The entries on the main diagonal are between 8 and 25 percentage points smaller than the corresponding direct estimates of the five-year transition matrix. This comparison suggests that the true persistence may be higher than estimated by first-order Markov chains.

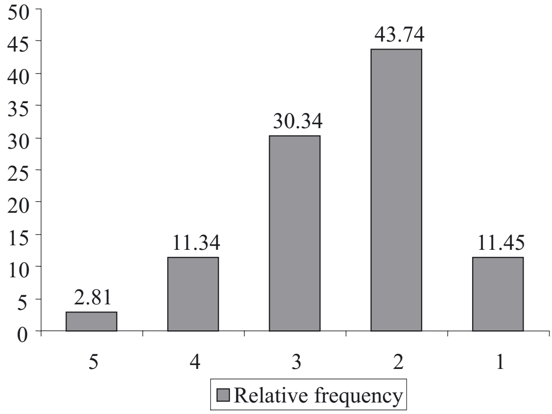


Figure 2: Health Distribution Across Self Assessed Health Categories Pooled for all Waves 1992 to 2005

The stationary distribution is observed after 25 iterations, with further changes in transition probabilities below 0.002 %. However, we can only claim stationarity relative to the underlying true state space if suitable measures of the inter-quintile range do not change over time. Since we apply fractile Markov chains based on quintile positions, the boundaries between states are not fixed over time, but are determined endogenously for each age-group. If the age-group-specific distribution becomes wider as cohorts of workers' age, this alone would reduce the number of observed transitions between health states, so that we may underestimate the true mobility unless we take changes in the entire distribution into account. To this end, the bottom section of Table 3 shows the development of the boundaries between the first and the second quintile as well as the range of quintiles two, three and four. All increase until cohorts reach the five-year age-group centred in 58 and decline thereafter, suggesting the variance of the distribution increases as workers age until they reach the typical age of retirement. Regressing the inter-quintile ranges on age using OLS for age-groups 28 to 58 reveals an average effect from moving to a higher age-group between 2.3 and 3.8 %.[20]

---

20 In addition to our earlier finding that the cross sectional measure of the health gradient becomes steeper as cohorts age, we find that persistence in the individual position on

Quantifying counterfactual scenarios. Figure 3 illustrates the empirical basis of our counterfactuals by showing the mean health status of early retirees and of workers in the highest income quintile by age-group. Younger early retirees tend to have particularly bad health. By contrast, the average worker in the highest quintile of the income distribution stays above the mean of the overall health distribution, indicated by 0 on the horizontal axis, until about age 50.

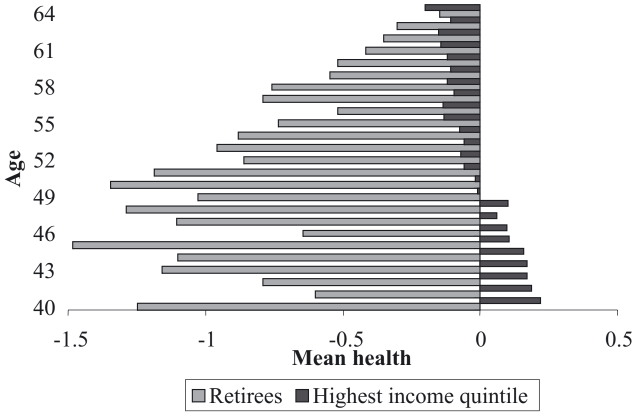


Figure 3: Mean Health at Retirement by Age

Based on our corridor of ex post predictions, Table 4 summarizes our two alternative counterfactual scenarios that are consistent with absence of the health gradient defined as no correlation between health and income. When – in our first counterfactual scenario – we assume equal access to existing healthcare and medical technology and all workers are able to maintain their health as individuals do in the highest quintile on the income distribution, each early retiree is assigned the age-specific mean health of the highest income quintile. In this scenario, retiring workers' health still declines with age, as can be seen by the age-dependency of the dark bars in Figure 3. But investment in better medical technology might lower the rate of health decline even below the rate currently observed in the highest income quintile. In our second counterfactual, we assume – as an example of the first-best – health does not decline with age at all so that all workers are able to maintain the state of health classified as very good on the ordinal scale in the SOEP.

In Table 4, the first panel compares our ex post predictions with the two counterfactual predictions for 40-year old workers, the second panel reports the potential gain from eliminating the health gradient and the third panel reports

the health gradient increases with age as well, when comparing estimates for the entire adult population with those above 40 years of age.

the potential gain from eliminating health decline. It appears that most working years are lost because workers' health deteriorates with age. Only between 0.32 and 0.63 working years would be saved per early retiree in our first, but between 2.63 and 3.04 years would be saved per early retiree in our second counterfactual.

Table 4

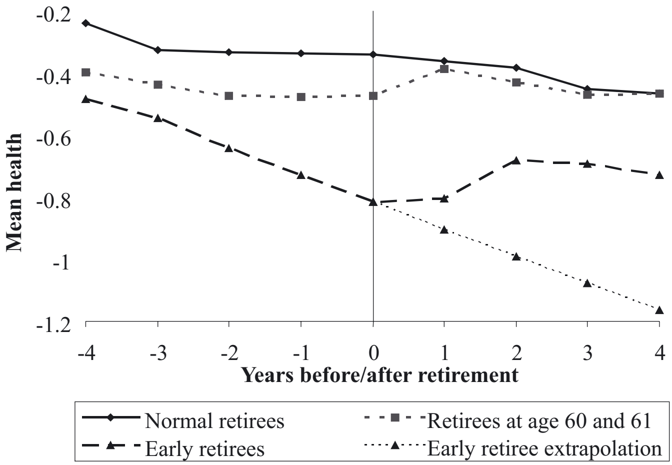
Lost Working Years due to the Health Gradient

	Covariate values at age 40 assumed constant thereafter	Covariate values at retirement assumed constant before
Mean time to retirement in days / Mean age at retirement		
Ex Post Predictions	6,947 / 59.02	6,791 / 58.59
Counterfactual mean health of upper income quintile	7,070 / 59.35	7,029 / 59.24
Counterfactual best health	7,953 / 61.79	7,953 / 61.79
Lost working years due to the health gradient		
total	4,020,458	7,816,393
per retiree	0.322	0.6264
months per retiree	3.86	7.51
Lost working years due to imperfect health		
total	32,900,000	38,000,000
per retiree	2.634	3.043
months per retiree	31.61	36.51

5. Quantifying the Social Costs

To complete our calibration of the social costs of health-related early retirement, we now quantify the health improvements that can be attributed to early retirement, relate these health improvements to subsequent mortality risk reductions and changes in survival probabilities and then briefly describe how we assign monetary valuations to these gains. To identify the relevant health improvements empirically, we would ideally like to compare each worker's health after retirement with a situation where that worker had not retired early. Aggregated across all early retirees, we would then have an estimate of the improvements in population health from early retirement. However, we do not observe the counterfactual and thus base our estimate on the nearest comparable scenario in the dataset, simplifying a methodology developed by Börsch-Supan / Jürges (2006). More specifically, we ask to what extent is the effect on health more favourable for those taking early retirement compared to those retiring at the normal retirement age. Consistent with non-health-related retirement pathways that become available after age 60, we divide the observations into two groups: all retirees who are younger than 60, but at least 40 years old at the moment of entry into retirement and the "normal" retirees defined as women above 60 years and men above 61, but in any case below 70 years of age.

Figure 4 shows the development of mean health of normal and early retirees after retirement. We can see that early retirees are on average a lot unhealthier than normal retirees, and controlling for age would reveal even larger differences. More importantly, the graph indicates that retirement has a positive effect on health only for early retirees. Average health of normal retirees continues to decline, whereas the average early retiree enjoys health improvements over the course of two years after retirement – and not just relative to the counterfactual extrapolation of the pre-retirement health trend, but also in absolute terms. Only in the third period after retirement does the health of early retirees start to decline again, but still at a lower rate than that observed for normal retirees.



Source: Own calculations based on the SOEP 1992–2005.

Figure 4: Development of Mean Health Around Retirement

We use this comparison of normal and early retirees to obtain an estimate of the effect of early retirement on health. To be conservative, we model the health improvement as a level effect, namely as the difference between the average health changes of normal and early retirees after two years. The effect is slightly larger for women than for men; women who retire early can improve their health measured on the cardinal health scale by 0.193 points on average, while men can improve their health by 0.181. We next assume the size of the effect to be linearly decreasing with age and obtain the slope of the linear relationship by comparing the average health improvements at the mean age of early retirees with retirees at age 60, the omitted age category, where we assume health changes due to retirement are zero. Since male early retirees are on average about a year older than female early retirees, the relationship between age and health improvements from early retirement is steeper for men.

Relating health improvements to mortality risk reductions. Bad health lowers the probability to survive at every age so that people who suffer from bad health die younger (Frijters et al., 2005 and Rehfeld, 2006). Using a Cox proportional hazard model, Jürges (2005) finds that the hazard to die is about 48 % higher for men in bad health compared to men in fair health, controlling for other relevant covariates recorded in the SOEP.

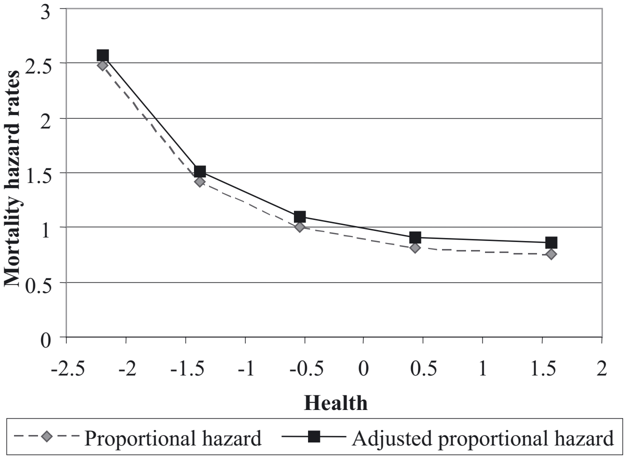
The estimated mortality hazards relative to those of men and women in fair health are 0.503* for men and 0.759 for women in very good health, 0.726** for men and 0.812 for women in good health, 1.481*** for men and 1.412*** for women in bad health, and 2.518*** for men and 2.479 *** for women in very bad health.[21] These estimates imply, for example, that a health improvement of 0.834 in our continuous measure reduces male workers' relative mortality hazard from bad health to the level of a worker in fair health. Assuming a piecewise linear relationship between health and mortality, a health improvement of 0.181 would reduce the hazard rate to die by 0.087. The magnitude of these effects is very similar in the case of women. To use this information in the calculation of social costs, we have to multiply the relative mortality hazards with a baseline hazard. Since Jürges' (2005) semi-parametric Cox model does not specify a baseline hazard function, we use standard life tables from the German federal statistics office to construct the baseline hazard to die. In doing so, we make three technical assumptions: First, the hazard ratios are applicable to the entire population although Jürges' (2005) analysis uses only the West German sub-sample. Second, for women in good and very good health the estimates of the proportional hazard to die are economically relevant although they are not significantly different from the baseline category in a statistical sense.[22] Third, we multiply the proportional hazard rates by a constant factor to make the average health status the baseline category as Jürges (2005) defined the middle health category as his reference category. Figure 5 shows that this results in a parallel upward shift of the proportional hazards dependent on health and that relative changes in health states are unaffected by this adjustment.

To obtain health-specific probabilities to survive, we multiply the proportional hazard to die conditional on health with the probability to die at a certain age and calculate the complement: $1 - \{\Pr(die|A, G, R) \times Hazard(die|H)\} = \Pr(s|A, G, R, H)$. The probabilities to survive conditional on age, gender and region are taken from the standard life tables of the years 2003/05 (Statistisches Bundesamt, 2006), which may lead to a slight underesti-

[21] The number of stars indicates the level of significance, 1, 5 or 10%.

[22] Including only the effect for women in bad health would lead to a downward bias of the health-state-specific probability to survive.

mation of age-specific mortality rates because life expectancy increased during the observed period. However, the difference between years 1992 and 2002 is significant only for males in Eastern Germany (Eisenmenger, 2005). To account for health-specific mortality after retirement, we make the additional assumption that mortality always increases with age. A worker who is in a bad state of health at retirement is assumed to *age* like an average person facing the same low survival probability at a much higher age.



Source: Own calculations based on the SOEP 1992 – 2005.

Figure 5: Mortality Hazard Rates Dependent on Health

Willingness-to-pay for mortality risk reductions. If retirement leads to a mortality reduction for early retirees, what should this be worth to society in monetary terms? Estimates of the value of a statistical life (VSL) can be derived from tradeoffs between labour income and job-related fatality risks and range between 0.5 and 16 million US dollar for a statistical life in the US – with a median of 5 million and an income elasticity of around 0.5 to 0.6 (Viscusi, 1993). A reasonable value for Germany will be lower than most US estimates. However, the literature for Germany is surprisingly scarce. Miller (2000) estimates a value of € 3.2 million for a statistical life in Germany. The European Commission (2000) suggests an upper limit of € 2.5 million for a VSL in environmental cost benefit analysis. Krewitt / Friedrich (2000) note that VSL estimates tend to be higher in medical than in environmental settings and suggest € 3 million as a reasonable value for the EU 15 countries. We apply this value and ignore the possibility that the value of a statistical life may vary with age.[23]

[23] The direction and size of the relationship between age and VSL is still debated. Jones-Lee (1989) finds an inverted-U-shaped relationship between VSL and age. Johan-

We also ignore historical evidence, such as Costa/Kahn (2004), that holding individual age constant, the value of a statistical life rises in real terms with an income-elasticity of about 1.6 over time – implying a significantly higher VSL today than estimated 10 or 20 years ago.

Social cost estimates. All costs are counted in the year that the individual enters retirement. We first calculate the present value of all future costs and benefits of early retirement for each individual. The first additive term in Equation 1 represents the present value of the lost labour income from the entry into early retirement until the pensioner reaches the statutory retirement age.[24] This involves double discounting – reflecting workers' time preference and the conditional probability to survive, which is below 100%. The second term includes the discounted gains from future mortality risk reductions, in terms of willingness-to-pay, between the age of entry into early retirement and age 100. Since we define early retirees as between 40 and 60 years of age, mortality risk reductions from early retirement can only be assigned to this group of retirees. In the last step, the social costs accrued for an individual are multiplied with the individual-specific person expansion factor and aggregated across all individuals to obtain the social costs at the population level. Here the actual expansion factors, as observed at the moment of retirement, not as in the first year of observation, are used so that the results can be interpreted for population totals.

The three panels in Table 5 report our results for the aggregated social costs and its two main components for the entire 13 year period in euro prices of the year 2001 – first on the basis of ex post predictions, second on the basis of counterfactual health states, and third on the basis of perfect health for all, the first-best counterfactual. Whereas the first column reports the values for the entire sample population, the five subsequent columns report those for workers in different quintile positions relative to the health gradient. In each panel, the entries in the first row report discounted labour income losses, the second row the gains from mortality reductions and the third row the social costs, calculated as the difference between the first two. Our ex post predictions suggest a total of about € 1,040 billion in discounted labour income losses and gains of about € 271 billion from reduced mortality risks, equivalent to 90,333 statistical lives

nesson/Johansson (1997) find that the value of a statistical life increases with age whereas several more recent studies, such as Krupnick et al. (2002), find that the VSL falls with age, especially after the age of 70. Using labour income data from the SOEP for the year 2000, Felder (2005) estimates that the value of remaining lifetime in Germany declines from approximately € 2.5 million at age 40 to € 1 million at age 60.

[24] Some retirees in our dataset, such as home makers or unemployed workers, have no labour income in the year prior to retirement. Since their income is assumed constant for the remainder of the potential working life, we implicitly assume that these retirees do not incur any social costs. This is consistent with our definition of social costs in Section II where we assume the value of early retirees' gain in leisure equals their lost contribution to GDP. This obviously applies to all cases of early retirement that leave leisure and GDP unchanged, such as the transition from unemployment into early retirement.

valued at € 3 million each. Defined as the balance between these two aggregates, the social costs are estimated at about € 769 billion, which amounts to an average of € 50 to 60 billion per year. The social costs attributable to workers in the lower quintiles (4 and 5) appear to be much larger than those in the higher quintiles (1 and 2), but these differences shrink in the two counterfactual scenarios.

Table 5

The Social Costs of Health-related Early Retirement, in billion euro

	Total	Health gradient				
		Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Ex post predictions						
Aggregated labour income losses	1,040	122	192	194	219	313
Health benefits from early retirement*	271	66.4	53	39	43	70
Social costs	769	55.6	139	155	176	243
Counterfactual health						
Aggregated labour income losses	995	98.6	171	184	207	334
Health benefits from early retirement*	230	21	25	38	48	98
Social costs	765	77.6	146	146	159	236
Value of direct mortality reductions from counterfactual health improvement	22,500	8,980	5,970	4,580	2,180	808
First-best counterfactual						
Aggregated labour income losses	600	74.1	111	110	110	195
Health benefits from early retirement*	0	0	0	0	0	0
Social costs	600	74.1	111	110	110	195
Value of direct mortality reductions from counterfactual health improvement	58,300	16,300	13,000	11,800	9,070	8,380

* Monetary value of mortality risk reductions implied by health improvements after early retirement.

The value of labour income losses in the counterfactual scenarios is influenced by two opposite effects: On the one hand, workers in better health tend to retire later so that labour income losses will be lower over the lifecycle. On the other hand, healthier workers also live longer which raises the probability that labour income is lost in the later years should a shock to health hit. In both counterfactuals, the first effect seems to dominate. By contrast, the value of health improvements from early retirement is lower the better a worker's health state. Our counterfactual assumption that all workers enjoy the mean health of the highest income quintile is responsible for the lower aggregate health benefits reported in the middle row of the second panel of Table 5. In the first-best counterfactual, the aggregate health benefits are zero because all workers are assumed to already be in the highest health state.

As workers directly value the higher levels of health that they enjoy in the two counterfactual scenarios, the final rows in the second and third panels of Table 5 report the estimated aggregate willingness-to-pay for the implied mortality reductions in the respective scenarios, as described at the end of sections II and IV. These estimates suggest that the potential welfare gains from better medical technology and more equitable and effective healthcare are likely to be several orders of magnitude greater than the estimated health benefits from early retirement in the current set-up.

Figure 6 shows the development of the social costs and its two main components over time. Substantial fluctuations are evident and the first trough coincides with the recession year of 1993. However, after that year the main components of social costs do not seem to be strongly correlated with the business cycle, suggesting that labour demand factors do not influence our results much. The peak of labour income losses lies in the year 2001, at about € 100 billion, almost twice the level observed in 2004, when labour income losses were at their lowest. Over the whole 13 year period, the social costs of health-related early retirement seem to be in decline, and the relatively low social costs observed since the pension law reforms in 2001 are mainly responsible for this trend.

6. Conclusion

A number of recent empirical studies have addressed the problem of the social costs of early retirement in Germany, yet identifying the part that is health-related has remained an unresolved task. Solving it requires *first* to identify the workers who retire on health grounds, even as eligibility criteria for disability pensions and the relative attractiveness and accessibility of alternative pathways to retirement, described in Viebrok (2003), change over time, and *second* to value the full costs and benefits in monetary terms, including the costs and benefits of changes in health. In this paper, we use a measure of self-assessed health for identification, emphasize the role of workers' stratification in health and income and estimate social costs that amount to almost € 60 billion in every year of the sample period, 1992 to 2005. Assuming constant individual productivity over time, the discounted aggregate losses of labour income streams are valued at about € 80 billion per year.

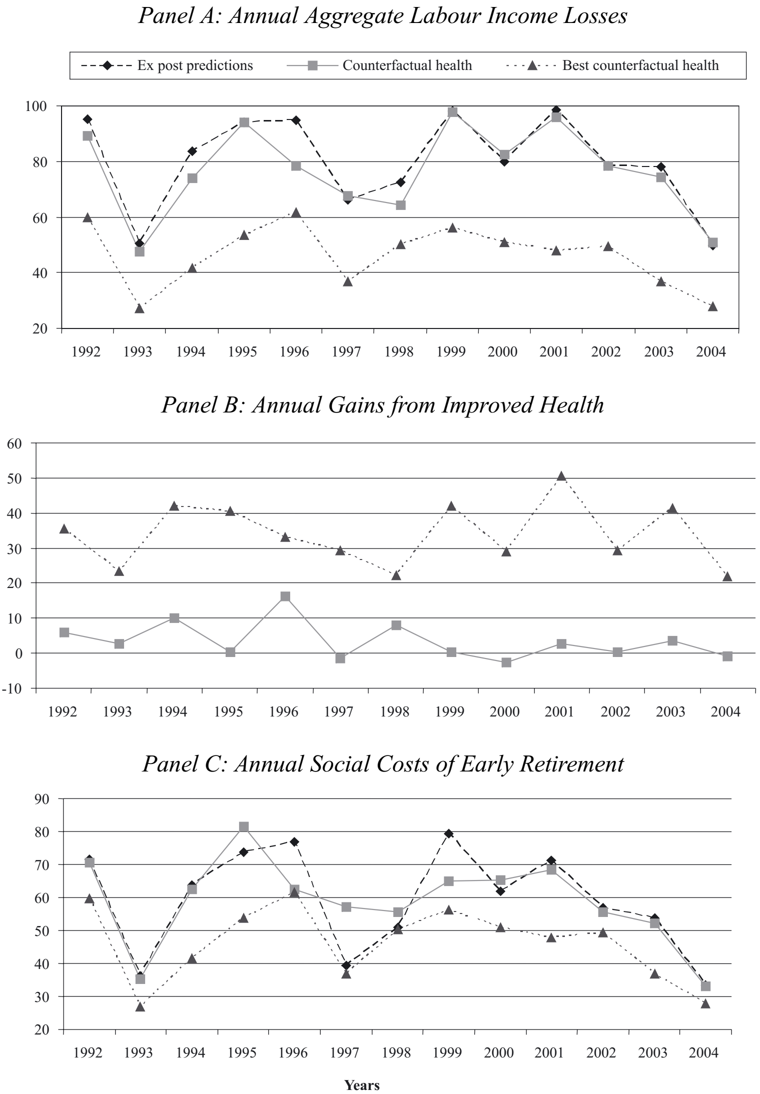


Figure 6: Components of Social Costs in Early Retirement, in billion euro

The social costs are lower than the labour income losses because early retirement can slow down and partly reverse work-related declines in health. We observe health improvements within two to three years after early retirement that cannot be observed for workers retiring at the "normal" age. These health benefits and the implied mortality risk reduction seem to be especially important for workers in the lowest quintile relative to their age-group-specific health gradient. To quantify the potential welfare gains from targeted investments in health and medical technology, we calibrate two counterfactual scenarios in which the health gradient is absent – first by assigning the mean health observed in the highest quintile and second by assigning the highest level of self assessed health to all workers at risk of early retirement. Our estimate of the social costs in 2003 alone is about twice as high as the liabilities in Germany's social budget that Rehfeld (2006) attributes to health-related retirement in 2003.

Our findings have important distributional implications for the design of health and pension policies that were absent in previous empirical studies of retirement timing based on structural models with a single representative agent. A particularly relevant implication is that pension policies with life expectancy links imposing an equal increase in the average age of retirement or an equal adjustment in pension levels on all quintiles of workers may fail to achieve the intended reduction of social costs. Workers in the bottom quintile tend to enjoy particularly large gains in life expectancy from the mortality reduction associated with early retirement as it tends to set in at a relatively young age. To maintain a given level of advanced life expectancy at some reference age, such as 40, poorer and less healthy workers would henceforth have to spend even more years in early retirement, *unless* the introduction of a higher statutory retirement age is accompanied not only by improvements in working conditions to reduce work-related health hazards, but also by substantial public health investments that target younger workers, defer their work-related health declines and raise labour productivity in the years up to and beyond the current statutory retirement age. Far beyond existing rehabilitation programmes, these investments would have to improve many aspects of acute and preventive care, especially for workers at risk of or already afflicted by chronic conditions.

As an additional payoff to society, such investments would also make raising the average retirement age more acceptable in terms of equity as they would help reduce the disproportionate financial burden that falls on workers with the lowest advanced life expectancy. In this vein, Hupfeld (2009) finds robust evidence of redistribution from the least able to the most able in Germany's statutory pension system, in line with earlier estimates of a positive correlation between the size of pension claims and individual longevity (Lauterbach et al., 2006; Himmelreicher et al., 2008). Without appropriate health investments, the inequities from pension policies that are actuarially fair or neutral for the average population are likely to increase as population aging appears to be driven mainly by gains in longevity among better educated and healthier people. Health investments for the poor should be seen as a complement to greater investments in their education and life-long learning as these facilitate the adoption of healthy lifestyles and jobs and increase the payoff from longer working lives.

References

Adams, H. P./*Hurd,* M. D./*McFadden,* D./*Merrill,* A./*Ribeiro,* T. (2003): Healthy, Wealthy and Wise? Tests for Direct Causal Pathways between Health and Socioeconomic Status, Journal of Econometrics 112, 3–56.

Becker, G. S. (1993): A Treatise on the Family, Enlarged Edition, Cambridge, MA.

Becker, G. S. (1973): A Theory of Marriage: Part I, Journal of Political Economy 81(4), 813–846.

Becker, G. S. (1974):. A Theory of Marriage: Part II, Journal of Political Economy 82(2, pt. 2), S11–S26.

Becker, G. S. (2007): Health as Human Capital: Synthesis and Extensions, Oxford Economic Papers 59, 379–410.

Börsch-Supan, A./*Jürges,* H. (2006): Early Retirement, Social Security and Well-being in Germany, National Bureau of Economic Research Working Paper No. 12303.

Bommier, A./*Stecklov,* G. (2002): Defining Health Inequality: Why Rawls Succeeds Where Social Welfare Theory Fails, Journal of Health Economics 21, 497–513.

Case, A./*Deaton,* A. (2005): Broken Down by Work and Sex: How Our Health Declines, in: D.A. Wise (ed.), Analyses in the Economics of Aging, Chicago/London.

Case, A./*Lubotsky,* D./*Paxson,* C. (2002):. Economic Status and Health in Childhood: The Origins of the Gradient, American Economic Review 92, 1308–1334.

Chotikapanich, D./*Creedy,* J./*Hopkins,* S. (2003): Income and Health Concentration in Australia, The Economic Record 79(246), 297–305.

Costa, D. L./*Kahn,* M. E. (2004): Changes in the Value of Life, 1940–1980, Journal of Risk and Uncertainty 29(2), 149–180.

Currie, J./*Stabile,* M. (2003): Socioeconomic Status and Child Health: Why is the Relationship Stronger for Older Children? American Economic Review 94(4), 473–503.

Deaton, A. S. (2001): Inequalities in Income and Inequalities in Health, in F. Welsh (ed.), The Causes and Consequences of Increasing Inequality, Chicago/London.

Deaton, A. S. (2002): Policy Implications of the Gradient of Health and Wealth, Health Affairs 21(2), 13–30.

Deaton, A. S./*Paxson,* C. (1998): Aging and Inequality in Income and Health, American Economic Review 88(2), 248–253.

Deutsche Rentenversicherung (2007): Statistische Analysen Referat 0631. http://www.deutsche-rentenversicherung.de/nn_15142/SharedDocs/de/Inhalt/04_Formulare_Publikationen/03_publikationen/Statistiken/Broschueren/ergebnisse_auf_einen_blick,templateId=raw,property=publicationFile.pdf/ergebnisse_auf_einen_blick.

DIW Berlin (2007): Overview SOEP, retrieved from: http://www.diw.de/english/soep/soepoverview/-27908.html.

Dwyer, D. S./*Mitchell,* O. S. (1999): Health Problems as Determinants of Retirement: Are Self-rated Measures Endogenous? Journal of Health Economics 18, 174–193.

Eisenmenger, M. (2005): Sterbetafel 2001/2003, Statistisches Bundesamt Wiesbaden. http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Con-tent/Publikationen/Querschnittsveroeffentlichungen/WirtschaftStatistik/Bevoelkerung/Sterbetafel,property=file.pdf.

European Commission (2002): Workshop on the Value of Reducing the Risk of Ill-Health or a Fatal Illness, Proceedings November 13, 2000; Brussels, Belgium. http://ec.europa.eu/environment/enveco/others/value_of_life.htm.

Felder, S. (2005): The Value of Life and the Value of New Medical Technology: Calculations for Germany, Mimeo. http://www.ifw-kiel.de/konfer/esf-ifw/newtech05/felder.pdf.

Finkelstein, A./*Luttmer,* E. F. P./*Notowidigdo,* M. J. (2008): What Good is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption, NBER Working Paper No. 14089.

French, E. (2005): The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour, Review of Economic Studies 72, 395–427.

Frick, J. R./*Grabka,* M. M. (2000): Personelle Einkommensverteilung und der Einfluss von Imputed Rent, DIW Discussion Paper No. 225. Retrieved from: http://www.diw.de/deutsch/publikationen/diskussionspapiere/docs/papers/dp225.pdf.

Frijters, P./*Haisken-DeNew,* J./*Shields,* M. (2005): Socio-economic Status, Health Shocks, Life Satisfaction and Mortality: Evidence from an Increasing Mixed Proportional Hazard Model, The Australian National University Centre for Economic Policy Research Discussion Paper No. 496.

von Gaudecker, H.-M./*Scholz,* R. D. (2007): Differential Mortality by Lifetime Earnings in Germany, Demographic Research 17, 83–108, Article 4.

Grossman, M. (1972): On the Concept of Health Capital and the Demand for Health, Journal of the Political Economy 80(2), 223–255.

Gustman, A./*Steinmeier,* T. (1986): A Structural Retirement Model, Econometrica 54 (3), 555–584.

Hagen, C./*Himmelreicher,* R./*Kemptner,* D./*Lampert,* T. (2011): Soziale Ungleichheit und Risiken der Erwerbsminderung, WSI-Mitteilungen 2011(5), 336–344.

Haisken-DeNew, J. P./*Frick,* J. R. (2005): Desktop Companion to the German Socio-Economic Panel (SOEP) Version 8.0.

Halliday, T. (2008): Heterogeneity, State Dependence and Health, The Econometrics Journal 11(3), 499–516.

Heiss, F./*Börsch-Supan,* A./*Hurd,* M./*Wise,* D. (2008): Pathways to Disability: Predicting Health Trajectories, in D. Cutler / D. Wise (eds.), Health at Older Ages, Chicago.

Herbertsson, T./*Orszag,* J. M. (2003): The Early Retirement Burden: Assessing the Social Costs of Continued Prevalence of Early Retirement in OECD countries, IZA Discussion Papers No. 816.

Himmelreicher, R./*Sewöster,* D./*Scholz,* R./*Schulz,* A. (2008): Die fernere Lebenserwartung von deutschen Rentnern und Pensionären im Vergleich, WSI-Mitteilungen 2008(5), 274–280.

Hostenkamp, G. / *Stolpe,* M. (2006): The Heath Gradient and Early Retirement: Evidence from the German Socio-economic Panel, Kiel Working Paper 1305.

Hostenkamp, G. / *Stolpe,* M. (2008): Optimal Health and Retirement Policy amid Population Aging, Kiel Working Paper 1428.

Hupfeld, S. (2009): Rich and Healthy – Better than Poor and Sick? An Empirical Analysis of Income, Health and the Duration of the Pension Benefit Spell, Journal of Health Economics 28, 427 – 443.

Johannesson, M. / *Johansson,* P.-O. (1997): Quality of Life and the WTP for an Increased Life Expectancy at an Advanced Age, Journal of Public Economics 65, 219 – 228.

Jones-Lee, M, W. (1989): The Economics of Safety and Physical Risk. Oxford.

Jürges, H. (2005): Self-assessed Health, Reference Levels and Mortality, Applied Economics 40, 569 – 582.

Kakwani, N. A. / *Wagstaff,* A. / *Van Doorslaer,* E. (1997): Socioeconomic Inequalities in Health: Measurement Computation and Statistical Inference, Journal of Econometrics 77, 87 – 103.

Krewitt, W. / *Friedrich,* R. (2000): Valuing the Environment in Monetary Units: Experience from the ExternE Project, SETAC World Congress, Brighton, 21.-25.05.2000, SETAC (Society of Ecotoxicology and Chemistry), Brussels.

Krupnick, A. / *Alberini,* A. / *Cropper,* M. / *Simon,* N. / *O'Brien,* B. / *Goeree,* R. / *Heintzelman,* M. (2002): Age, Health and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario, Journal of Risk and Uncertainty 24 (2), 161 – 186.

Larsen, M. / *Gupta,* N. D. (2004): The Impact of Health on Individual Retirement Plans: A Panel Analysis Comparing Self-reported Versus Diagnostic Measures, Aarhus School of Business Department of Economics Working Paper 04-7.

Lauterbach, K. / *Lüngen,* M. / *Stollenwerk,* B. / *Gerber,* A. / *Klever-Deichert,* G. (2006): Zum Zusammenhang zwischen Einkommen und Lebenserwartung, Studien zu Gesundheit, Medizin und Gesellschaft 1, 1 – 13.

Miller, T. R. (2000): Variations between Countries in Values of Statistical Life, Journal of Transport Economics and Policy 34(2), 169 – 188.

Muurinen, J.-M. / *Le Grand,* J. (1985): The Economic Analysis of Inequalities in Health, Social Science and Medicine 20(10), 1029 – 1035.

Osaki, S. (1992): Applied Stochastic Modelling, Berlin.

Rehfeld, U. G. (2006): Gesundheitsbedingte Frühverrentung, Gesundheitsberichterstattung des Bundes, Heft 30, Berlin.

Siddiqui, S. (1997): The Impact of Health on Retirement Behaviour: Empirical Evidence from West Germany, Health Economics 6(4), 425 – 438.

Shiller, R. J. (2003): The New Financial Order: Risk in the 21st Century, Princeton.

Shkolnikov, V. M. / *Scholz,* R. / *Jdanov,* D. A. / *Stegmann,* M. / *von Gaudecker,* H.-M. (2007): Length of Life and the Pensions of Five Million Retired German Men, European Journal of Public Health, 1 – 6.

Statistisches Bundesamt (2006): Sterbetafeln 2003/2005. http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Content/Statistiken/Bevoelkerung/Geburten Sterbefaelle/Tabellen/Content100/SterbetafelFBNL,property=file.xls.

Van Doorslaer, E./*Koolmann,* X. (2004): Explaining Differences in Income Related Health Inequalities Across Europe, Health Economics 13, 609–628.

Vermeulen, F./*Kalwij,* A. (2006): Labour Force Participation of the Elderly in Europe: The Importance of Being Healthy, Revised Version, CentER Discussion Paper No. 2006-87.

Viebrok, H. (2003): Disability Pensions in Germany, in C. Prinz (ed.), European Disability Pension Policies. Ashgate.

Viscusi, W. K. (1993): The Value of Risk to Life and Health, Journal of Economic Literature 31, 1912–1946.

Whitehouse, E. (2007): Life-expectancy Risk and Pensions: Who Bears the Burden? OECD Social, Employment and Migration Working Papers No. 60, Paris.

Appendix

The German Socio-economic Panel (SOEP) is a longitudinal microeconomic dataset on a wide variety of topics that is designed to be representative for the non-institutionalized resident population in Germany. The SOEP dataset was provided by the Deutsches Institut für Wirtschaftsforschung (DIW Berlin). It started in 1984 and has been repeated annually since then. In 1990 it was expanded to include the East German population. In order to ensure a large sample size, additional samples were drawn over the course of time to include immigrants and to offset ongoing panel attrition. In addition the SOEP has been adjusted and modified in a number of ways to ensure international comparability, for example through the rephrasing of questions and the introduction of new topics (DIW, 2007). Foreign residents, as well as East German residents and high income households are over-sampled in the SOEP, while elderly people above 70 are under-represented. To adjust for disproportionate sampling of subgroups, non-response and attrition in the course of time the SOEP provides pre-calculated cross sectional and longitudinal weighting factors. In each cross section the weighting factors are normalized according to population marginals in the micro census and add up to the non-institutionalized resident population in Germany. Longitudinal weighting factors are provided in the form of staying probabilities to account for panel attrition (Haisken-DeNew/Frick, 2005). The average life tables for the years 2003/05 were provided by the Statistisches Bundesamt. All regression analyses were performed using the statistical software package Stata 8.0. Some of the preparative calculations for the social costs formula were done in Excel version 2003.