The Kiel Institute for the World Economy
Duesternbrooker Weg 120
24105 Kiel (Germany)

Kiel Working Paper No. 1305

The Health Gradient and Early Retirement: Evidence from the German Socio-economic Panel

by

Gisela Hostenkamp and Michael Stolpe

December 2006

The Health Gradient and Early Retirement: Evidence from the German Socio-economic Panel[ϒ]

Gisela Hostenkamp and Michael Stolpe

Abstract:

This paper examines the role of the health gradient – the positive correlation between household income and health – in individual retirement behavior, using data from the German Socio-economic Panel (GSOEP). We first estimate agegroup-specific health gradients and find their slope increases with age, but declines among retired workers. We then estimate a variety of parametric and semi-parametric duration models and find that workers' position relative to the agegroup-specific health gradient has about the same explanatory power as self-assessed health and income together. We argue our method promises better predictions of the long-term impact of policies affecting the health gradient on workers' timing of retirement amid population aging. Our findings also underline the importance of imperfect medical technology in reconciling the human capital theory of health demand with the observation of more rapid declines in health among less educated workers.

Keywords: Health gradient; Retirement behaviour; Duration analysis; Germany

JEL-Classification: H51, H55, I12, J26

Corresponding author:
Michael Stolpe
Kiel Institute for the World Economy
24100 Kiel
Germany
Telephone: +49–431–8814 246
E-mail: michael.stolpe@ifw-kiel.de

[ϒ] This paper is part of a research project on „The Changing Distribution of Household Health Spending in Germany“ within the Kiel Institute's Global Health Economy Program.

Contents

1 Motivation

The health gradient – the positive correlation between income or financial wealth and health – links the two major variables commonly thought to determine the timing of individual retirement from the labor force. But empirical evidence of a causal relationship between a worker's position on the health gradient and the propensity to retire early is surprisingly sparse. The present paper provides the first systematic investigation of this relationship in Germany. Using a large dataset from the German socio-economic panel (GSOEP) that covers the 1992 to 2004 period, we find that workers' position relative to the agegroup-specific health gradient has about the same explanatory power as self-assessed health status and income together and therefore serves as a near-perfect substitute for these conventional covariates in duration analyses of individual retirement behavior.

Recent empirical studies have shown that the health gradient in income or wealth increases with age during people's working lives. As health tends to deteriorate with age and – more pronounced – with working years, the net value of the expected income that can be achieved through continued work tends to decline as well. When declines in health are predominantly work-related, retiring often promises to slow down or pause the further deterioration of health. Early retirement is therefore often particularly attractive for manual workers whose health tends to suffer more from work-related wear and tear and who are vastly over-represented at the bottom end of the health gradient.

With work-related declines in health often irreversible and the decision to retire usually a once and for all decision, the positive correlation between health and income tends to exacerbate the stronger private retirement incentives for people with poor health compared with those in good health. Needless to say, the pension system also plays an important role and may further enhance the private value of early retirement if the early-retirement annuity reduction is less than actuarially fair and/or if the size of annuity payments rises less than proportionally with a person's prior contributions to the pension system. These are salient characteristics of Germany's public pension system. Börsch-Supan and Jürges (2006) provide details of recent reforms and a discussion of retirement incentives in Germany.

A better understanding of the relationship between workers' position on the health gradient and their individual propensity to retire is of particular interest when a country enters a long period of population aging. Under these conditions, early retirement becomes an increasing burden to society that needs to be addressed through strategic policy reforms on the basis of reliable long-term forecasts. Identifying a causal relationship between the health gradient and retirement behaviour does not only provide a new tool to forecast future retirement behavior on the basis of social and economic determinants of the health gradient. It can also establish a powerful new rationale for health and educational policies that target the health gradient. The problem of early retirement can then no longer be addressed by pension policies alone, but must be addressed by investments in health and other forms of human capital.

Smith (1999) surveys a vast empirical literature that has mainly sought to identify, describe and understand patterns of the health gradient across space and time, not primarily motivated by policy issues. A particularly interesting strand of this literature has tried to untangle the direction of causality between health and wealth, but has largely focused on subsets of the problem only. For example, Adams et al. (2003) use Granger causality tests that reveal causality from wealth to health among US pensioners above 70 years of age, whose retired status eliminates the possibility of reverse causality from health to the ability to work. But a related study by Michaud and Soest (2004) on US couples aged 51 to 61, using dynamic panel methods, finds no conclusive evidence of causality from wealth to health, only from health to wealth. In any case, these studies are of limited interest from a policy point of view as they do not attempt to explain how the health gradient is created in the course of people's lives.

To develop an informative life-cycle analysis, Deaton and Paxson (1998a and 1998b) use US cohort data and, in line with similar findings from other countries, observe an almost fourfold increase in the health gradient until about age 60, associated with a systematically increasing variance of self-reported health outcomes. In a similar vein, Case and Deaton (2005) establish empirical correlations that suggest a causal link between investments in education and the evolution of income and health over the life-cycle – consistent with an explicit intertemporal model of health demand in which a person's financial, educational and health capital serve as partial substitutes. In the main piece of evidence, manual workers' health is found to decline faster than that of non-manual workers.

Studies on retirement behavior in Germany, such as Börsch-Supan and Jürges (2006) and Siddiqui (1997), also point to health as the major empirical determinant. Methodologically, we depart from these studies in several ways. For a start, we use an ordinal measure of *self-rated health* on a five-point scale that the GSOEP reports since 1992 and that is comparable to subjective health measures widely used in recent international studies, such as Larsen and Gupta (2004) for Denmark and McGarry (2002) for the US. Previous authors on Germany have used a ten-point measure of *personal health satisfaction* which the GSOEP has recorded since 1984; its coefficient of correlation with self-rated health status in our sample is -0.72. A multi-country comparison of subjective health measures is provided in Börsch-Supan et al. (2005).

Our interest in the health gradient relates mainly to the role of health policy and medical technology in reducing the association of persistent health inequalities with socio-economic status and in averting a rising social burden from early retirement as populations age. However, by focusing on the health gradient, we also hope to help overcome the pervasive methodological problems of justification bias and endogeneity in previous retirement studies. A justification bias can arise when early retirees report bad health as a socially acceptable excuse, which may lead to an overestimation of the health impact on retirement timing relative to other time-varying covariates, such as income. Larsen and Gupta (2004) provide a detailed empirical investigation of this problem. By encompassing the two major time-varying risk-factors of early retirement, which we hypothesize to be negatively correlated with each other for any given person in the short term when additional income can be generated at the expense of health, the health gradient reduces the possible scope for a purely health-related estimation bias. Moreover, unlike surveys that focus exclusively on retirement behaviour, the GSOEP does not ask for the reason of retirement, nor does it ask for self-rated health and retirement status at the same time so that respondents to the health question are unlikely to make a conscious link to their retirement decision.

On the other hand, endogeneity may lead to an underestimation of the health impact on retirement timing when early retirement is motivated by the prospect of halting work-related declines in health, a possibility consistent with evidence in Börsch-Supan and Jürges (2006). By disallowing separate influences of health and income, our measure of workers' position relative to the agegroup-specific health gradient essentially allows bad health to have a stronger impact when the gradient is steeper and there is little additional income from continued work to offset the pull of bad health into retirement. By allowing health to play a larger role when it matters more, the potential estimation bias from endogenous health is reduced.

The remainder of this paper proceeds as follows. Section 2 discusses the relevant theory and derives the hypotheses we wish to test. Section 3 describes our dataset from the GSOEP. Section 4 explains our empirical methods and presents the findings of our duration analyses. Section 5 discusses these findings in the context of prior literature, outlines potential policy implications and identifies promising directions for further research. Section 6 concludes.

2 Theory and Hypotheses

2.1 Theory

The literature offers few theoretical models that can shed light on the evolution of the health gradient over the life cycle and serve as a source of testable hypotheses with implications for the policy issues we seek to address. The most important strand of this literature builds on Grossman's (1972) human capital model of health demand. This model is famous for introducing the idea that a person's health can be interpreted as a form of capital from which "health services" flow that are consumed over time. The stock of health is an exogenous endowment at birth, but an endogenous variable forever after. It tends to decline with age until the person dies. The rate of depreciation is assumed to vary with age and to be influenced by the intensity with which the stock of health is used for work or consumption as well as by the availability and quality of medical care to repair health shocks.

Muurinen and LeGrand (1985) adapt Grossman's human capital model to study the evolution of health inequalities across social classes in Britain. They assume less educated people to work predominantly in manual jobs that make greater direct demands on health, carry greater risks of health shocks, such as injuries, and let the stock of health depreciate faster than non-manual jobs. The better educated a worker, the better her chances of finding a non-manual job. Matching in the labour market thus imposes constraints on people with less human and financial capital and forces them to make more intensive use of their health capital in earning a living. Lack of financial capital may prevent people from making educational investments or from setting up as an entrepreneur. In consequence, less educated and poorer people are predicted to experience a faster deterioration of their health with age.

We, too, consider educational and health capital to be defining components of human capital and follow Grossman's (1972) and Muurinen and LeGrand's (1985) emphasis on the partial substitutability of three basic forms of capital – financial capital, either inherited or accumulated through savings from labour income, educational capital, obtained primarily through schooling, and health capital, maintained primarily by physical exercise, dietary restraint and medical care. We note in passing that, while existing stocks of capital may be used as partial substitutes in generating earnings, the phase of accumulation may rather be characterized by a complementarity of investments in health and education in the build-up of human capital. For example, in the years before a young person first enters the labour force, complementarity can arise either from different levels of health at birth, which then trigger differential rates of educational investment, or from different levels of inherited financial wealth, which set different incentives to invest in both health and schooling. The following focuses on the implications of substitutability between education and health in generating income and therefore takes people's different initial assets, health and educational capital as given.

We build on Case and Deaton (2005) who invoke the human capital model to motivate an empirical analysis of the evolution of health inequality over the life-cycle. They seek to explain the puzzling observation from many countries that health worsens with age more rapidly among poorer people, who also tend to start their working lives with lower health. Since this is directly related to our own motivation, we briefly summarize their main line of argument using Figures 1 and 2 as an illustration.

If the working poor are mainly found in manual jobs, why should these workers – in addition to low wages – pay the price of ruining their health? This is indeed a puzzle for economic theory. As Case and Deaton (2005) point out, with medical technology that fully repairs health shocks, the standard interpretation of the human capital model of health demand cannot account for a positive correlation between the rate of health depreciation and the demands on health imposed by a person's chosen work. The standard interpretation instead predicts that people purchase medical care to offset any work-related health deterioration and keep the marginal utility of their health stock equal to its user costs, assumed constant over time. In Figure 1, perfect medical technology is depicted by the health repair function with parameter $\hat{\theta}$. It guarantees that any person starting with H^0 in period t can purchase the appropriate amount of medical care to maintain or restore health to H^0 in period $t+1$, regardless of the rate of health depreciation δ in period t. The model's dynamic equilibrium, as Case and Deaton (2005) show, implies that only the rate at which *aging may increase* the rate of work-related health depreciation will eventually lead to greater declines in health; but the rate of increase will generally be *lower* among manual workers than among those with higher education, since manual workers' aging encounters a higher age-adjusted rate of work-related health depreciation to begin with.

Case and Deaton (2005) argue that the observation of more rapid declines in health among manual workers can be reconciled in a plausible way with the human capital model by assuming that the technology of medical care is *less* than perfect. The level of physical deterioration from work will then determine the rate of decline in health because the marginal utility of health cannot be maintained equal to its user cost, as first-best optimality would require if medical technology were perfect.

Figures 1 and 2 summarize these insights in an intuitive way by illustrating the optimality conditions that Case and Deaton (2005) identify formally: Consumers maximize life-cycle utility $U = \sum_0^T (1+\rho)^t \, \nu\left(c_t\right)$, where ρ is the rate of time preference, T the exogenous length of life and $\nu\left(.\right)$ is the instantaneous utility from consumption c_t, subject to the present-value budget constraint $\sum_0^T (1+r)^{-t} c_t + \sum_0^T (1+r)^{-t} p_m \, m_t = A_0 + \sum_0^T (1+r)^{-t} y_t \left(H_t, z_t\right)$, where r is the market rate of interest, p_m the price and m_t the chosen quantity of medical care, A_0 denotes initial assets and $y_t\left(H_t, z_t\right)$ earnings, assumed to increase with the stock of health, H_t, and with the physical effort in manual labor, z_t. The stock of health evolves according to $H_{t+1} = \theta m_t + \left(1 - \delta_{t,z}\right) H_t$, where θ is the efficiency with which medical care repairs health damages, and $\delta_{t,z}$ is the rate at which health deteriorates in period t, given the physical work effort z_t.

Case and Deaton (2005) derive the first-order conditions for consumption and health, $\nu_{ct} = \lambda \left(1+\rho\right)^t \left(1+r\right)^{-t}$ and $\frac{p_m}{\theta}\left(r + \delta_{t,z}\right) = y_{ht}$, where subscripts with respect to h and c denote partial derivatives. The Lagrange multiplier, λ, can be interpreted as the shadow price of lifetime wealth and is constant over the life-cycle. Figure 2 illustrates these conditions for alternative values of z, using the concept of marginal efficiency of health investments. The downward-sloping *MEC*-curve shows the marginal product of health in terms of earnings for different health stocks and can be interpreted as the demand schedule for health capital, the horizontal line *S* as the supply-curve. With diminishing marginal productivity of health in generating earnings, the first-order conditions imply that at any point in time and throughout life, the health stock will be higher (i) the lower the price of medical care, p_m, i.e. the smaller is α in the third quadrant of Figure 1, and the greater medical efficiency in restoring health, θ, i.e. the greater the level of H_{t+1} that can be reached in the second quadrant of Figure 1 before the health repair function levels off into its vertical section, (ii) the lower the rate of health deterioration, $\delta_{t,z}$, i.e. the steeper the ray in quadrant one, and the weaker the effect of diminishing returns to health in earnings, i.e. the less concave the earning function in quadrant four, (iii) the lower the rate of time preference, ρ, (iv) the larger lifetime income effects, that reduce λ, such as higher initial assets and initial health and the lower prices of general consumption goods over the lifetime.

In the special case of equality between the rate of interest and the rate of time preference, health will decline over the life-cycle if and only if the rate of health deterioration, $\delta_{t,z}$, increases with age. Changes in the stock of health ΔH_{t+1} are linked to the physical deterioration of health by the identity $\Delta H_{t+1} = \theta\, m_t - \delta_{t,z} H_t$. With perfect medical technology, all health damages can be repaired fully so that death is entirely voluntary and people will choose a finite life only if the rate of health deterioration increases with age. In Figure 1, perfect medical technology is depicted by $\hat{\theta}$; and the associated medical expenditure in the third quadrant will be higher the greater the health damage to be repaired.

When manual workers choose to improve their earnings at the expense of a faster deterioration in health, moving from $y(H|\, z = 0)$ to $y(H|\, z = 1)$ in the third quadrant of Figure 1, they are in effect selling part of their health capital. The relevant first-order condition requires that the marginal gain in earnings from additional manual work is equal to the marginal health costs, the product of health stock and the marginal effect on user costs: $\frac{\partial y_t}{\partial z} = \frac{p_m}{\theta} \cdot \frac{\partial \delta_{t,z}}{\partial z} \cdot H_t$. At a fixed level of education, those with more initial health are less willing to undertake heavy labor to improve their earnings because they would incur a greater loss from an increase in the rate of health depreciation. If the health stock is optimally adjusted to its user costs, those with better education will simply enjoy a greater health stock throughout life because – unlike manual workers – they gain little or no additional income by making a greater physical work effort, such as $z = 1$ in Figure 1.

In the absence of perfect medical technology, people cannot maintain their health stock equal to user costs and work-related wear and tear will lead to an actual decline in health. Workers whose lack of education and wealth forces them into manual jobs are then predicted to experience higher rates of health deterioration and to see their health decline more rapidly with age. In Figure 1, imperfect medical technology is depicted by the health repair function labeled $\tilde{\theta}$ that levels off into its vertical section at point *B*.

Viewed from a life-cycle perspective, the opportunity to retire early may represent an important mechanism to constrain the impact of declining health on total life-time income – by limiting the work-related deterioration of retirees' health and by giving them a chance to receive annuity payments from pay-as-you-go pension systems for a longer period of time. The private value of early retirement will be enhanced if the early-retirement penalty in the size of annuity payments is less than actuarially fair and/or if the size of annuity payments rise less than proportionally with a person's prior contributions to the pension system. These are salient characteristics of Germany's current public pension system.

From a social point of view, the opportunity to retire early under Germany's present rules and regulations may limit the *private* incentives of those with relatively poor health and earnings prospects to invest in the development and maintenance of their personal human capital. Early retirement is hence unlikely to be a *socially* efficient way of reducing the health gradient. A more efficient policy strategy, especially in aging populations, would target the health of younger cohorts in order to prevent the health gradient from becoming steeper as these cohorts age.

2.2 Hypotheses

Against this background, our empirical research questions and hypotheses can be summarized as follows: (*i*) Does a person's health status affect the timing of retirement in Germany? We hypothesize: the lower a person's health status, the earlier will he or she retire. (*ii*) Does the health gradient increase with age? We hypothesize: the older the agegroup from which the sample is drawn, the steeper will be the health gradient in income, at least among active workers whose age-related health decline varies with the type of their work. (*iii*) Is the individual position relative to the agegroup-specific health gradient correlated with the timing of retirement? We hypothesize that people with both low income and bad health will tend to retire earlier than those where only one of these conditions holds.

The concept of the health gradient imposes a specific constraint on the way health is confounded by income in the timing of retirement. Health is allowed to have a stronger impact, the steeper the health gradient. When health is strongly negatively correlated with income, the pull of bad health into retirement is much less offset by high income from continued work. In Figure 1, this would mean a greater variance of workers' individual positions along the horizontal axis that indicates their age-specific health stocks. In Figure 2, the age-specific *MEC*-curve shifts to the left for those who have already suffered a series of persistent health shocks and irreversible declines in their health. By imposing this constraint on our empirical estimates, we hope to improve the predictability of retirement behavior in the future as we believe that predictions based on tested theories are more reliable in the long term than those based on more correlations in the available data.

3 Data

The German Socio-economic Panel (GSOEP) is a microeconomic dataset on a wide variety of topics that is based on annual surveys designed to be representative of the population living in Germany. It started with the first annual survey for the western part of Germany in 1984. After the fall of the Berlin wall, it was expanded to include people in both parts of Germany from 1990 onwards. Over time, the GSOEP has been adjusted and modified in a number of ways, such as the rephrasing of questions and the introduction of new topics, to ensure international comparability in line with the objective to form the European Community Household Panel. For a detailed description of the genesis and development of the GSOEP, see the introduction and references provided at http://www.diw.de/english/sop/uebersicht/ index.html.

Because health is not the main focus of the GSOEP, the choice of meaningful variables for longitudinal analyses that cover all years since 1984 is limited. As a measure of individual health, for example, previous studies, such as Börsch-Supan et al. (2004), Frijters et al. (2005) and Siddiqui (1997), have relied on the variable *personal health satisfaction*, a measure of self-assessed health that is reported on a scale from 1 to 10 and may be less comparable with subjective health measures in most countries.

To make our analysis internationally comparable, we take advantage of the variable *self-rated health-status* (SRHS) which the GSOEP has reported since 1992. It is measured on an ordinal scale from 1 to 5, with one representing the best and five the worst individual health status, and included in our regression analyses untransformed, thus ignoring the possibility of applying the “empirical normal transformation” suggested by van Praag and Baarsma (2001); Börsch-Supan and Jürges (2006) find that this transformation actually makes little difference when applied to the variable personal health satisfaction. We believe that for our purpose, self-rated health status is superior to more objective, yet often incomplete, measures of health because respondents’ subjective self assessment of overall health will include both physical and mental aspects. The latter is inherently difficult to assess objectively, but increasingly important for a variety of reasons, including long-term changes in the workplace.

To identify the *timing* of individual retirement, the GSOEP offers several variables to choose from, depending on the specific focus of the analysis. Studies on the implications of retirement for government spending have often used the year of first pension payment. For our focus on the timing of the retirement decision, we require more exact data on the time of withdrawal from the workforce. We therefore use the month when a person first declares himor herself retired to define the retirement event. Unlike other retirement variables in the GSOEP, this information is available on a monthly basis, albeit only after the completion of each calendar year. Although we may loose some observations when respondents drop out of the GSOEP within that year, we prefer our more accurate measure over the annual indicator for the receipt of retirement payments as this also includes orphans and widows, which are not of interest in our context. By using the GSOEP waves from 1992 to 2004, our analysis avoids structural time series issues that may be associated with the event of German unification. In addition, we confine our sample to adults above 40 years of age who had not yet retired until the beginning of the sampling period and estimate only single-spell duration models because we understand retirement as an absorbing state, which is much less likely to hold for workers retiring before the age of 40.

To construct a robust measure of the health gradient in income, we start with the natural logarithm of household income after tax and government transfers, henceforth post-government income, which we believe to best capture the combined disposable yearly income of all household members together. We do not limit this to wage income because entrepreneurial income generation, too, may require some effort related to health. After transforming all prices in Euros and applying the OECD consumer price deflator, our measure of household income is expressed in prices of 2001. Total household income has been used in related studies, such as the analysis of mortality determinants by Frijters et al. (2005). In addition, we run regressions in which household income is adjusted for the number of household members, first including children and second including only adults. Per capita household income has been widely used to define socio-economic status, for example recently by Bender and Habermalz (2005).

Before we turn to our evidence on the determinants of retirement, we wish to clarify whether there is a meaningful health gradient in Germany and how it might change in the course of people's lifetime. To begin with some stylised facts about the relationship between income and health, Figure 3 plots self-rated health status for five-year age-intervals, comprising the life-cycle from entry age into the labor force to death, for strata defined by quartiles of pre-government household income, in Panel A, and post-government household income, in Panel B. Along the horizontal axis, we plot the five-yearly age intervals for which the strata-specific average self-rated health is calculated. In line with the prediction of a rising health gradient, the lower the income-quartile, the faster the deterioration of health. The self-rated health of those in the highest income group falls only gently until about 60 years of age, but thereafter even faster than the population-average so that self-rated health is equalized at around 80 years of age. The process of health deterioration in all other income-quartiles appears to pause or at least slow down markedly between the age of 55 and 70, which may be related to these groups' relatively high frequency of early retirements that take away the work-related risks to health.

Another look at these stylized facts is provided in Figure 4 where we use the results from linear regressions of self-rated health on household income, reported in Table 9, to illustrate the rising slope of the agegroup-specific health gradient. The legend gives the central age in each five-year interval defining the agegroups. The rising slope appears to be caused almost entirely by greater health declines among lower income groups, until people reach about 60 years of age. Thereafter, the slope of the health gradient declines. This decline, which appears as a rotation of the regression curve around a fixed mid-point, can probably be attributed at least partly to a rising death toll among those with lower self-rated health, a process of self-selection out of the sample which is also bound to be responsible for part of the slowdown in health decline observed among the lower income groups in Figure 3. Table 9 reveals that the agegroup in which the slope of the health gradient reaches its maximum varies between the early 50s and the late 60s depending on the exact definition of income used in the calculation.

Our measure of workers' position relative to the health gradient is obtained by the orthogonal projection of each worker's individual location in the agegroup-specific health-income diagram onto the log income axis; we then subtract the agegroup-specific median value to eliminate general aging-related changes in health or income. As an alternative measure of workers' position, we also construct a set of dummy variables, one for each quintile of workers in ascending order of their relative position on the health gradient, so that we can control for an increasing dispersion in the continuous health gradient measure as cohorts age. The quintiles are defined by five intervals with equal shares of the sample population in the respective agegroup.

Figure 4 nicely illustrates the dominant influence of income on our definition of a worker's relative position when the health gradient is flat and of health when the health gradient is steep. The steeper the health gradient, the more is the income effect augmented by a correlated health effect.

Using nonparametric duration analysis for workers aged 40 and older, Figure 5 reveals that the correlation between workers' quintile position and the hazard of early retirement is indeed increasing with age. Analysis time, plotted on the horizontal axis, is measured in days beginning at age 40. By comparing Panel A, based on household income per adult in calculating workers' position on the health gradient, and Panel B, based on self-rated health strata in our sample, we can see that the role of health dominates the health gradient in explaining early retirement in younger agegroups, but looses its dominance in older agegroups. Note that the five-point scale for self-related health is now defined from 0 to 4, as stated in the appendix; this ensures that the lowest self-assessed state of health, not some hypothetical zero-state further down, defines the baseline hazard. The ordinal health scale from 0 to 4 is therefore used in all our duration analyses.

The non-health related legal opportunities for early retirement after age 60 are clearly reflected in the striking drop of workers' "survival" rates around analysis time 7300. This drop is evident in all quintiles of the health gradient and also in all self-rated health strata, except for the lowest. This reinforces the conclusion that purely health-related cases of early retirement have pre-empted other potential motives among those with the lowest health status.

Our duration analyses further include the number of years of education, ranging from 7 to 18, to measure a person's educational attainment beyond the legally required minimum and a dummy variable indicating whether a person is a public sector employee, to account for differences in the pension systems for public and private sector employees. Other control variables that previous studies concerning the health gradient or retirement decisions found to be significant and which we therefore include as well are dummies for gender, residence location in Eastern or Western Germany, employment status, and marital status. Table 1 in the appendix has more information on some of these variables.

4 Empirical Analysis

4.1 Methods

We believe the best strategy to understand the relationship between the health gradient and the timing of early retirement in Germany and test our hypotheses is to use modern duration analysis, also known as survival analysis in medical research, where the underlying question is quite simply: How long does a person “survive” as a worker and what determines observed differences in retirement timing?

We hope to use estimates from stratified survival analyses in future research to help predict the likely impact of selected reform proposals for health and pension policies on the evolution of retirement behaviour and returns from human capital formation in Germany's aging population. We therefore estimate parametric survival models, based on the Weibull distribution assuming monotone hazard rates for early retirement. Only with parametric estimates can we make out-of-sample predictions of workers' time to retirement, conditional on the covariates included in our analysis. We use semi-parametric estimates of Cox proportional hazard models merely as a point of comparison to assess the appropriateness of the Weibull distribution in our context. Finally, we also estimate shared frailty and stratified survival models with strata defined on the basis of workers' position relative to the agegroup-specific health gradient, which the preceding section identified by means of linear regression analyses.

Duration models have been a standard choice of method in the literature on the determinants of mortality and early retirement for some time, although a number of problems with this methodology are well known. Studies attempting to establish the direction of causality between health and early retirement are plagued by problems of endogeneity and unobservable individual heterogeneity. Recent studies have tried various novel strategies to overcome these problems. For example, Roberts et al. (2005) estimate the impact of health on the retirement hazard, controlling for confounding factors such as income and pension entitlement, by tracking the same individuals over time. By constructing a health stock variable, they also attempt to overcome the justification bias and remove measurement errors that are often associated with subjective health measures.

A strategy to account for unobserved heterogeneity that increases over time is developed by Frijters et al. (2005) who use data from the GSOEP to study the role of socio-economic characteristics and personal health satisfaction in determining the length of life. Unobserved persistent health-related shocks that may lead to cumulative health deteriorations are introduced through an individual time-varying unobservable that is defined to have expectation 0 over the whole population and follows a Wiener process, also known as a random walk or unit root process. The authors thus hope to overcome the inherent distortion from an unduly large emphasis on unobserved health differences at the beginning of analysis time in the standard mixed proportional hazard model that pays no attention to the potentially more important accumulation of health deteriorations during the analysis time.

While these issues are surely relevant in our context, we also see problems in the approach of Frijters et al. (2005), such as the issue of identification when individuals who are hit by severe health shocks have a greater hazard of dying, resulting in a selective stock of individuals at risk of retiring thereafter. Moreover, it is not clear how the approach could be used to make out-of-sample predictions that we feel should be based on tested hypotheses about causal relationships. More specifically, we do not see how the approach of Frijters et al. (2005) could be used to test whether the persistence of individual health shocks is related to socio-economic determinants, such as education, occupation and the level of income or financial wealth. But these determinants are at the center of discussions about causality, and causality must be established to make longer term predictions and discuss policy issues.

We will therefore account for heterogeneity by estimating *stratified* duration models where the baseline hazard is allowed to differ between population strata defined on the basis of workers' position relative to the agegroup-specific health gradient, but constrained to be the same within each of these strata. This allows for time-varying heterogeneity due to unobserved persistent and cumulative health shocks *between*, but not *within* strata. We believe that this approach allows us to perform meaningful tests of the implications of the human capital model of health demand and to make the kind of out-of-sample predictions of time-to-retirement that are required to assess the merits of specific government policies.

As an alternative hypothesis, we also consider *shared* frailty models, where we do not believe that individual heterogeneity is the same for all people in our sample, but that it is related to workers' individual position relative to the agegroup-specific health gradient. We therefore estimate shared frailty models in which groups are defined on the basis of workers' position relative to the agegroup-specific health gradient and compare these with unshared frailty models.

A problem that we ignore is the potential selection bias that may arise when those with particularly bad health are prevented from retiring early by an early death. This bias may reduce our estimates of the impact of health and the health gradient measure to some extent; and we plan to correct for this in future research.

4.2 Regression Estimates

Tables 2 to 8 report a series of duration regressions with which we seek to demonstrate how our preferred specification performs relative to models that do not stratify. Analysis time is always set to begin at each worker's biographical age of 40 years. In most of these tables, we report parametric survival regressions, assuming a Weibull distribution, in the proportional hazard metric as well as semi-parametric Cox proportional hazard models, with the same set of covariates, in order to assess the appropriateness of the Weibull distribution for our hypotheses tests and predictions. The more similar the estimated hazard ratios in these regressions turn out, the less constraining – and hence the more appropriate – must be the choice of a Weibull distribution in our parametric models.

In most tables, regressions are reported in pairs, using the same method, so that we can directly compare the explanatory power of the two covariates income and self-reported health status separately with the explanatory power of the single covariate measuring workers' individual position relative to the agegroup-specific health gradient.

In addition, we include the following covariates as control variables in the regressions: household wealth, measured by the deflated imputed rental value of owner-occupied houses; employment status, measured by a dummy variable equal to one in case of no employment; a further dummy equal to one for public sector employees; the number of years in formal education in addition to a seven-year minimum time of schooling; a dummy variable equal to one for females; a dummy variable equal to one for residence in Eastern Germany; a dummy variable for marital status with the value zero for all married persons and one for all unmarried persons, whether they are bachelors, divorced or widowed. We also include each person's year of birth as a covariate because we wish to account for the possibility that time-varying circumstances in childhood and upbringing that are reflected in the year of birth influence people's health in later life and their propensity to retire early.

Tables 2 to 4 provide results on basically one and the same model that neither stratifies, nor allows for unobserved heterogeneity, but is re-estimated for different definitions of the relevant income variable, included both as a direct covariate and as an input in our health gradient measure. The first two columns in these tables report results from Weibull regressions, the second two from corresponding Cox proportional hazard models with the same covariates, and the last two report Weibull regressions with robust standard errors, obtained by clustering on workers' quintile position on the health gradient.

Our principal findings are robust across the different specifications. Table 2 starts with our simplest model in which each household's total income is used. Self-rated health status has a highly significant positive effect: a one-point worsening on the five-point SRHS scale increases the hazard rate of retirement by approximately 40 percent. Household income also has a highly significant effect, albeit in the opposite direction: a unit increase in the log of total household income lowers the hazard rate by 20 percent. In columns 2, 4 and 6, workers' individual position relative to the age-specific health gradient has about the same explanatory power as income and self-reported health status, when these are not included as covariates, and a one unit- improvement lowers the retirement hazard by 33 percent. Table 2 also suggests that neither a person's employment or marital status, nor residence in former East Germany have a significant impact on the hazard rate. But the year of birth matters: A delay of one year apparently results in a drop of the hazard rate by ten percent, according to our Weibull estimates.

Table 3 and 4 report the corresponding results of regressions using two alternative definitions of per capita household income, first household income divided by the number of all household members and second household income divided by the number of adult household members. These adjustments hardly change our estimates for most covariates, while both measures of per capita household income are actually insignificant in the Weibull model, shown in columns 1 and 5 of Tables 3 and 4, respectively. Income per adult household member is significant in the Cox regression and a unit increase of the log measure lowers the hazard rate by 15 percent. The impact of our health gradient measure is lower when household income is adjusted for adults and children, but almost unchanged from the regressions with total income when income is adjusted for adult household members only. Also the estimated hazard ratios for the other covariates show little change, but the dummy for marital status becomes a significant covariate in the Weibull regressions without robust standard errors; those who are unmarried appear to have an almost 20 percent higher hazard of retirement.

As for other covariates we employ as controls, a birth date delayed by one year shows up in a ten percent lower hazard in the Weibull and a six percent lower hazard in the Cox model.

Every additional unit in our logarithmic measure of wealth increases the retirement hazard rate by roughly two percent. Employment in the public sector increases the hazard rate by 35 to 45 percent whereas additional years of education lower the hazard rate by up to 5 percent per year. The hazard rate of women is more than 10 percent higher than that of men. The dummy variable for marital status turns out to be not significant in almost all specifications.

Table 5 reports the results of duration analyses in which our continuous measure of workers' position on the health gradient is first supplemented and then replaced by a set of dummy variables that indicate workers' quintile position relative to their agegroup-specific health gradient. As explained in section 3, the set of quintile position dummies offer a control for the possibility of time-varying variance in the continuous health gradient measure of workers' relative position. Employing the set of quintile dummies along with continuous health-related covariates, as reported in columns 1 and 2, allows us to perform a nested test of the health gradient impact versus separate covariates for self-rated health and income. It turns out that including the additional dummies does not improve the quality of the Weibull model, which we report in columns 1 and 2, relative to the corresponding regressions in Table 2; the log likelihoods barely change and the implied values of the Akaike information criterion (AIC), which equals $\mathrm{A} = -2\log L + 2k$, where L is the likelihood and k the number of all estimated parameters in the model, are virtually the same.

Columns 3 to 5 report regressions without the continuous health-related covariates so that the dummies for quintile position carry all the information about workers' position on the health gradient. We have three columns because we again distinguish between our three income concepts in the definition of quintile positions on the health gradient. We find in all three specifications that workers in the lowest quintile have the highest hazard of retirement and that the hazard rates tend to be lower the higher the quintile to which a worker belongs. For example, moving from the lowest to the second lowest quintile reduces the hazard by one quarter when the health gradient is calculated on the basis of total household income, by one third when it is calculated on the basis of household income per capita and by 40 percent when it is calculated on the basis of income per adult household member. Moving to the highest quintile reduces the hazard rate by almost one half, albeit slightly less when the health gradient is based on per capita income including children.

Overall, both the log likelihoods and the implied values for the Akaike information criterion suggest that using income per adult household member when calculating quintile positions on the health gradients results in the best model. In fact, this model represent a small improvement over the corresponding regression with our continuous health gradient measure, reported in columns 2 and 6 of Table 4. By contrast, the regressions with quintile positions based on total household income and per capita household income, reported in columns 3 and 4 of Table 5, perform slightly worse than the corresponding regressions in Table 2 and 3.

Table 6 reports the results of shared frailty models, an approach to time-invariant unobserved heterogeneity that is conceptually similar to random-effects models in panel analysis. We let our observations cluster in each quintile position on the health gradient and assume that there is intra-cluster correlation, but no correlation between clusters. Essentially, we assume that observations are correlated within those quintile groups due to some overall group characteristic, a frailty in statistical jargon, that is not being measured. Provided this assumption is correct, the shared frailty model will produce robust standard errors, which reflect inter-cluster variability of parameter estimates, and therefore yield more accurate standard errors and test statistics for the impact of covariates on the retirement hazard.

Our results show that the estimated hazard ratios of the covariates are largely similar, but the frailty variance, theta, is significantly different from zero only in columns 5 and 6, where income per adult household member is entered as a covariate and as an input into the health gradient measure, respectively. In line with this, only Model 5 and 6 show some very modest improvement in terms of log likelihood and the Akaike information criterion when compared to the corresponding models of the preceding tables. As a point of comparison, we have also estimated unshared frailty models (available on request), but in these, the null hypothesis of a gamma frailty is rejected in all models that use any of the three variants of our income measure to calculate the health gradient. In combination, we take these findings as some evidence of clustering in workers' quintile positions on the health gradient and of unobserved heterogeneity between strata that may be consistent with correlated health shocks within strata and path dependence in the accumulation of strata-specific health shocks over time.

Table 7 reports the results of stratified duration models, again comparing the Weibull, in Panel A, and the Cox specifications, in Panel B. In stratified models, the baseline hazard is allowed to differ for different groups, or strata. We stratify for people's individual quintile position relative to the agegroup-specific health gradient and use the lowest quintile as the base category. In parametric stratified estimation, each group is allowed to have its own baseline hazard function, but these hazard functions are constrained to be of the same family, in our case the Weibull distribution. In stratified Cox estimation, the assumption of the same baseline hazard for everyone, multiplied by their relative hazard, is relaxed in favor of a different baseline hazard for each group, but the coefficients on the covariates are constrained to be the same across groups.

We find that stratification improves the overall quality of our regressions markedly, as indicated by the log likelihood and the implied values for the Akaike information criterion. Moreover, all the dummy variables for quintile position turn out to be highly significant covariates with the expected sign. Among the Weibull models, those based on income per adult household member, reported in columns 5 and 6 of Table 7a, perform best. All stratified Weibull regressions confirm that the hazard of early retirement is highest in the lowest quintile on the health gradient, and at its lowest in the fourth and fifth quintile. In Model 5, the estimated effect of moving from the lowest to the second lowest and from the third to the fourth quintile are particularly strong. In addition, a unit-improvement in the continuous health gradient measure is estimated to reduce the retirement hazard of 29 percent. But in Model 6, household income per adult is no longer a significant covariate. In fact, only total household income, used in Model 2, retain a significant impact in these stratified regressions.

The Cox regressions reported in Table 7b for models with the continuous health gradient measure among the covariates are broadly in line with the Weibull estimates, although wealth is no longer significant. All stratified regressions have greater log likelihoods and clearly lower implied values for the Akaike information criterion than the corresponding models of the preceding tables.

Table 8 reports results of stratified duration regressions in accelerated time metric, also known as accelerated failure time (AFT). In these models, stratification does not only allow the shape of the baseline survival function to vary across groups, but also allows time to accelerate or decelerate in different patterns for individual strata. This class of duration models thus allows more flexibility than the simple stratified models reported in Table 7. It is conceptually very close to the idea that workers in one and the same strata share more or less the same intensity of persistent health shocks that become a cumulative effect on the retirement hazard.

We wish to emphasize that the estimated coefficients from an accelerated time model must be interpreted differently because the effect of the covariates is now to accelerate time by a factor of $exp\left(-\beta_x x_j\right)$. These estimates are not directly comparable to those obtained in any of the preceding regressions. All five AFT models reported in Table 8 stratify with respect to workers' quintile position on the agegroup-specific health gradient, but Model 1 and 5 do not include the continuous health gradient measure as a covariate. Instead, they include self-rated health status and household income, the latter adjusted for the number of adults in Model 5. As in the models of Table 7, the lowest quintile position on the health gradient is used as the base category. The results show that both the scale and the shape of the hazard is significantly different from the base category in each of the other strata. In terms of overall quality, however, these regressions do not appear to improve, or only a little bit, on the corresponding models reported in Table 7.

An important advantage of parametric duration models is that one can use their results to predict hazard and conditional survival functions as well as the mean or median time to retirement. The predicted mean time, for example, is calculated as the integral of the survival function from zero to infinity, given each observation's covariate values and the estimated model parameters. Such predictions of the time to retirement can be used in future research to calibrate welfare effects of policies that influence the health gradient and other significant determinants of the time to retirement in Germany.

Figure 6 provides a first visual impression of selected predictions of the hazard rates after age 40 for each quintile of workers' relative positions on the health gradient. Panel A shows these predictions for married and working females born in 1950 with ten years of education. Panel B displays the corresponding predictions for men. The differences between the quintile groups are very similar for men and women: Those in the lowest quintile have a fast-rising hazard rate initially, but some deceleration thereafter. Those in the next two quintiles see an almost linearly increasing hazard rate, whereas those in the two highest quintiles face a very small hazard initially that subsequently seems to rise exponentially, converging with the hazard rates of quintiles two and three.

These hazard rates are predicted on the basis of our estimates from Model 5 in Table 7a. We choose Model 5 because it has the lowest Akaike information criterion of all regressions that include our measure of workers' position on the health gradient as a covariate. To be sure, the Akaike information criterion of this model could be further improved by including self-rated health and household income per adult member as separate covariates (see Model 6) instead of only indirectly through the health gradient. However, we set out to make predictions on the basis of the health gradient because we believe that this may become an important policy target as population aging proceeds and accelerates.

5 Discussion

While health inequality has become an important issue on the European policy agenda in recent years (see, for example, Judge et al. 2005), research on the health gradient in income, wealth and other indicators of socio-economic status remains a controversial topic in economics. Indeed, only limited progress has been made on the methodologically difficult issue of causality; various pieces of evidence still seem to contradict each other, as in Adams et al. (2003) and Michaud and van Soest (2004). Against this background, our empirical exploration into the explanatory power of the health gradient for the timing of retirement in Germany suggests a new research strategy. We believe the health gradient could become a valuable paradigm in the creation of a consistent framework to think about health in aging societies and a focusing device for empirical research on health-related investments. This may lead to the development of new tools to predict the long-term consequences of health-related social policies and identify efficient policy targets in anticipation of population aging.

Putting the health gradient and retirement in context is not merely a question of equity, but of paramount importance for economic efficiency. The timing of retirement has pervasive efficiency implication for the individual, for social security and for the economy as a whole: For the individual, alternative retirement ages imply specific combinations of present and future amounts of income and leisure, with considerable influence on economic well-being as well as health status. For social security, the timing of retirement determines when a person turns from a contributor of taxes to a recipient of benefits, so that regulations affecting this timing are particularly important in public pension systems based on defined benefits. For the macro-economy, even small variations in the average retirement age can result in large changes in the aggregate supply of labor, with potentially important implications for productivity and economic growth.

Much recent empirical research, surveyed in Lindeboom (2006), points to poor health as one of the most important reasons for early retirement, often before financial incentives and wealth effects. The association between various measures of health and wealth or income has been established in different populations and across the whole range of the income distribution. Although cross-country evidence is ambiguous on the hypothesis that the health gradient is steeper the more unequal the general access to health care, Deaton (2002) suggests that quicker access of better educated and wealthier individuals to the newest and often most potent medical technologies may be an important explanation for the persistence of the health gradient. We therefore believe that not merely the improvement of access, but also the development of new medical technologies that facilitate access to effective therapies for aging-related diseases, particularly those concentrated in the lower social strata, will be an important task for private and public investment in the future. A robust analytical framework to forecast the impact of such investments is sorely needed.

Our paper has begun to explore the merits of using the health gradient, instead of health and income separately, to explain the timing of retirement. We point out that this comes at a cost since we are in effect constraining two of the most significant regression coefficients known from the exiting literature. Is this price worth paying? The advantages we hope to obtain from imposing the constraint are in terms of prediction and policy analysis and therefore may ultimately be verifiable only with hindsight. But we are convinced that a model-based analysis is needed to better understand causality, improve forecasts and conditional predictions for alternative scenarios and thus make progress in designing and implementing efficient policy strategies. Past research in this field has been insufficiently concerned with testing the implications and predictions of explicit theoretical models.

To implement the research strategy we have outlined, the next steps in our work will concentrate on the issue of temporal persistence in workers' relative position on the health gradient, which we plan to study with the help of Markov models. We believe that Markov models can also be used to predict people's changing position on the health gradient in response to specific policy reforms, and that this may actually be easier and more promising than attempting to predict future states of health. Our research strategy will further address the question of social costs by calibrating intertemporal models based on predictions from our empirical studies of retirement behavior. We hope the findings from this research program will prove useful in assessing the extent to which Germany and other OECD countries should raise their official retirement age in response and anticipation to future progress in medical technology that will help to cure diseases faster and more fully than in the past.

Our present findings already underline the importance of imperfect medical technology in reconciling the human capital model of health demand with the observation of greater declines in health in less educated workers. We therefore disagree with Deaton's (2002) view that targeting the health gradient is not an appropriate mission for government policy. He may be right when policy instruments are limited to the redistribution of existing resources. But we think that targeted investments in medical technology will play a key role in creating new resources that are sorely needed to reduce health inequalities and address the health gradient in Europe's aging populations of the 21st century.

6 Conclusion

In this paper, we have estimated semi-parametric and parametric survival models, the latter based on the Weibull distribution assuming monotone hazard rates, to explain the timing of early retirement in Germany. We have also estimated shared frailty and stratified survival models with strata defined on the basis of workers' position relative to the agegroup-specific health gradient, which we have identified by means of linear regression analyses.

We conclude that our new measure of workers' position relative to the agegroup-specific health gradient has about the same explanatory power as self-assessed health status and income together and therefore serves as a near-perfect substitute for these conventional covariates in survival analyses of individual retirement behaviour. When we have used workers' quintile position relative to their agegroup-specific health gradient as a categorial covariate, our results appear to be mainly driven by workers in the lowest quintile, who tend to retire at approximately twice the hazard rate of those in the highest quintile.

We believe that our research can be easily extended to perform more specific stratified survival analyses, test hypotheses about cohort- and strata-specific retirement behaviour and make predictions about the likely long-term impact of government policies that affect the health gradient and workers' mobility relative to the health gradient during their lifecycle. By corroborating some of the basic predictions of the human capital model of health demand with imperfect health repair, our findings underline the importance of improvements in medical technology and access in overcoming the threat of an increasing social burden from a steepening of the population-wide health gradient amid population aging.

References

Adams, P., M. Hurd, D. McFadden, A. Merrill, and T. Ribeiro (2003). Healthy, Wealthy, and Wise? Tests for Direct Causal Paths between Health and Socioeconomic Status. *Journal of Econometrics* 112: 3–53.

Bender, K. and S. Habermalz (2005). Are there differences in the Health-Socioeconomic Status Relationship over the Life Cycle? Evidence from Germany. IZA Discussion Paper No. 1560.

Börsch-Supan, A., K. Hank and H. Jürges (2005). A New Comprehensive and International View on Ageing: The Survey of Health, Ageing and Retirement in Europe. Discussion Paper Series, Mannheim Research Institute for the Economics of Aging, No. 75–2005.

Börsch-Supan, A., and H. Jürges (2006). Early Retirement, Social Security and Well-being in Germany. NBER Working Paper 12303. Cambridge, Mass.: NBER.

Börsch-Supan A., R. Schnabel, S. Kohnz and G. Mastrobuoni (2004). Micro Modelling of Retirement Choices in Germany. In: J. Gruber and D. Wise (eds.), Social Security Programs and Retirement around the World. Micro-Estimations: Chicago, 285–343.

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

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

Deaton, A.S. and C. Paxson (1998a). Aging and Inequality in Income and Health. *American Economic Review Papers and Proceedings of the 110th AEA Annual Meeting* 88 (2): 248–253.

Deaton, A.S. and C. Paxson (1998b). Health, Income, and Inequality over the Life Cycle. In: D.A. Wise (ed.), Frontiers in the Economics of Aging. Chicago and London: The University of Chicago Press.

Frijters, P., J. Haisken-DeNew and M. Shields (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.

Grossman, M. (1972). The Demand for Health: A Theoretical and Empirical Investigation. Columbia University Press for the National Bureau of Economic Research: New York.

Jones A.M., E. van Doorslaer, T. Bago d'Uva, S. Balia, L. Gambin, C. Hernández Quevedo, X. Koolman and N. Rice (2006). Health and Wealth: Empirical Findings and Political Consequences. *Perspektiven der Wirtschaftspolitik* 7 (Special issue): 93–112.

Judge, K., S. Platt, C. Costonges and K. Jurczak (2005). Health Inequalities: A Challenge for Europe. An independent expert report commissioned by, and published under the auspices of, the UK Presidency of the EU.

Larsen, M., and N.D. Gupta (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.

Lindeboom, M. (2006). Health and Work in Older Workers. In: A.M. Jones (ed.), Elgar Companion to Health Economics. Edward Elgar: Cheltenham.

McGarry, K. (2002). Health and Retirement: Do Changes in Health Affect Retirement Expectations? NBER Working Paper 9317. Cambridge, Mass.: NBER.

Michaud, P.C., and A.H.O. van Soest (2004). Health and Wealth of Elderly Couples: Causality Tests Using Dynamic Panel Data Models. CentER Discussion Paper No. 2004–81.

Muurinen, J.M. (1982). Demand for Health: A Generalised Grossman Model. *Journal of Health Economics* 1: 5–28.

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

Roberts, J., N. Rice and A.M. Jones (2005). Health and Retirement in the UK: Duration Analysis of the British Household Panel Survey. Mimeo University of Sheffield.

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

Smith, J.P. (1999). Healthy Bodies and Thick Wallets: The Dual Relationship between Health and Economic Status. *Journal of Economic Perspectives* 13: 145–166.

Van Doorslaer, E. and A.M. Jones (2004). Income-Related Inequalities in Health and Health Care and the Aging European Population. *Health Economics* 13: 605–608.

Van Praag, B.M.S., and B.E. Baarsma (2001). The Shadow Price of Aircraft Noise Nuisance. Tinbergen Institute Discussion Paper 2001–010/3.

Appendix

I. Tables

Table 1: Summary Descriptive Statistics

Panel A: Data

Category	total	per subject:	mean	min	median	max
No of subjects	18759					
No of records	74699		4.0	1	3	12
Time at risk	26438174		1409.4	31	1095	4383
Retiring	6425		0.34	0	0	1

Panel B: Variables

Health	
Self-rated health status	Ordinal 5 point scale [0 = very good; 4 = very bad]
Education	
No of years of formal education	7 years = 0; More = 1;2;3…
Economic status	
Employment status	Working = 0; No work = 1
Public service employee	No = 0; Yes = 1
Income	
Post-government household income	HH labour income + HH asset income + Private transfers + Public transfers + Social Security pensions + HH Private retirement income – Total HH Taxes
Real household income	Post-government household income in Euro prices of 2001
Log household income	Log of Real household income
– per capita of all household members	Log of (Real household income / No of persons in household)
– per adult household member	Log of (Real household income / No of adults in household)
Wealth	
Wealth	Log of (Imputed owner-occupied house rental value in Euro prices of 2001)
Health gradient	
Relative position on the agegroup-specific health gradient	Continuous
Quintile position	Discrete [1=lowest; 5=highest], abbreviated in Figure 5, based on household income per adult, as rel_pos4.

Table 2: Results of Duration Regressions for Early Retirement, using Household Income

	1	2	3	4	5	6
Hazard Ratios	Weibull	Weibull	Cox	Cox	Weibull	Weibull
					robust standard errors	
Self-rated health	**1.36**		**1.40**		**1.36**	
z	9.30		10.09		5.24	
P>|z|	0.000		0.000		0.003	
Health gradient		**0.66**		**0.64**		**0.66**
z		-7.57		-8.16		-7.47
P>|z|		0.000		0.000		0.000
Household income	**0.80**		**0.82**		**0.80**	
z	-3.53		-3.26		-5.92	
P>|z|	0.000		0.001		0.000	
Wealth	**1.02**	**1.02**	**1.01**	**1.01**	**1.02**	**1.02**
z	2.62	2.83	0.83	1.09	2.99	3.25
P>|z|	0.009	0.005	0.406	0.276	0.003	0.001
No employment	**1.50**	**1.32**	**1.33**	**1.07**	**1.50**	**1.32**
z	1.27	0.85	0.88	0.20	2.60	1.56
P>|z|	0.205	0.395	0.376	0.839	0.009	0.118
Public sector	**1.37**	**1.34**	**1.46**	**1.46**	**1.37**	**1.38**
z	5.01	5.11	5.93	5.99	6.19	6.92
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
Years of education	**0.98**	**0.98**	**0.957**	**0.96**	**0.98**	**0.98**
z	-1.98	-1.59	-3.70	-3.11	-2.55	-2.52
P>|z|	0.047	0.112	0.000	0.002	0.011	0.012
Female	**1.13**	**1.13**	**1.19**	**1.18**	**1.13**	**1.13**
z	2.07	2.08	2.78	2.70	3.60	5.24
P>|z|	0.039	0.038	0.005	0.007	0.000	0.000
Eastern residence	**1.11**	**1.07**	**1.22**	**1.17**	**1.11**	**1.07**
z	1.51	0.89	2.83	2.23	1.76	1.26
P>|z|	0.130	0.371	0.005	0.026	0.078	0.209
Unmarried	**1.03**	**0.94**	**1.02**	**0.90**	**1.03**	**0.94**
z	0.31	-0.81	0.25	-1.36	0.23	-0.68
P>|z|	0.758	0.417	0.800	0.174	0.816	0.495
Year of birth	**0.89**	**0.89**	**0.94**	**0.94**	**0.89**	**0.89**
z	-14.58	-14.70	-8.10	-8.18	-8.74	-8.63
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
/ln_p	**1.04**	**1.08**			**1.04**	**1.08**
z	20.77	22.25			3.84	4.09
P>|z|	0.000	0.000			0.000	0.000
p	2.83	2.94			2.83	2.94
1/p	0.35	0.34			0.35	0.34
No. of subjects	10728	10728	10728	10728	10728	10728
No. of failures	1236	1236	1236	1236	1236	1236
Time at risk	16778048	16778048	16778048	16778048	16778048	16778048
Log likelihood	-822.68	-845.08	-8006.03	-8030.11	-822.68	-845.08
No. of obs.	47255	47255	47255	47255	47255	47255
LR chi^2	400.31 (10)	355.52 (9)	301.13 (10)	252.97 (9)		
Prob > chi^2	0.000	0.000	0.000	0.000		
Akaike information c.	1669.36	1712.16	16036.06	16082.22	1669.36	1712.16

Table 3: Results of Duration Regressions for Early Retirement, using Per Capita Household Income including Adults and Children

	1	2	3	4	5	6
Hazard Ratios	Weibull	Weibull	Cox	Cox	Weibull	Weibull
					robust standard errors	
Self-rated health	**1.36**		**1.40**		**1.36**	
z	9.41		10.15		5.68	
P>|z|	0.000		0.000		0.000	
Health gradient		**0.77**		**0.72**		**0.77**
z		-4.84		-6.17		-2.97
P>|z|		0.000		0.000		0.003
Household income	**0.92**		**0.89**		**0.92**	
z	-1.29		-1.82		-0.90	
P>|z|	0.197		0.069		0.366	
Wealth	**1.01**	**1.02**	**1.00**	**1.00**	**1.02**	**1.02**
z	2.24	2.20	0.55	0.54	1.43	1.37
P>|z|	0.025	0.028	0.584	0.591	0.152	0.170
No employment	**1.66**	**1.51**	**1.43**	**1.20**	**1.67**	**1.51**
z	1.59	1.28	1.11	0.57	3.42	2.48
P>|z|	0.113	0.200	0.268	0.565	0.001	0.013
Public sector	**1.37**	**1.39**	**1.46**	**1.48**	**1.37**	**1.39**
z	5.05	5.20	5.99	6.10	5.28	6.89
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
Years of education	**0.97**	**0.97**	**0.95**	**0.96**	**0.97**	**0.97**
z	-2.63	-2.33	-4.06	-3.57	-1.83	-1.94
P>|z|	0.009	0.020	0.000	0.000	0.067	0.053
Female	**1.15**	**1.18**	**1.21**	**1.23**	**1.15**	**1.18**
z	2.34	2.75	3.09	3.42	5.53	5.94
P>|z|	0.018	0.006	0.002	0.001	0.000	0.000
Eastern residence	**1.16**	**1.12**	**1.26**	**1.23**	**1.16**	**1.12**
z	2.09	1.68	3.25	2.92	1.77	1.41
P>|z|	0.037	0.093	0.001	0.003	0.077	0.159
Unmarried	**1.16**	**1.18**	**1.14**	**1.15**	**1.15**	**1.18**
z	1.86	2.10	1.68	1.78	1.26	1.56
P>|z|	0.063	0.036	0.094	0.075	0.207	0.119
Year of birth	**0.89**	**0.89**	**0.94**	**0.94**	**0.89**	**0.89**
z	-14.56	-14.54	-8.01	-7.91	-9.30	-9.89
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
/ln_p	**1.05**	**1.08**			**1.05**	**1.08**
z	21.13	22.29			4.49	4.80
P>|z|	0.000	0.000			0.000	0.000
p	2.87	2.95			2.87	2.95
1/p	0.35	0.34			0.35	0.34
No. of subjects	10728	10728	10728	10728	10728	10728
No. of failures	1236	1236	1236	1236	1236	1236
Time at risk	16778048	16778048	16778048	16778048	16778048	16778048
Log likelihood	-828.01	-861.18	-8009.65	-8043.48	-828.01	-861.18
No. of obs	47255	47255	47255	47255	47255	47255
LR chi^2	389.65 (10)	323.32 (9)	293.90 (10)	226.24 (9)		
Prob > chi^2	0.000	0.000	0.000	0.000		
Akaike information cr.	1680.02	1744.36	16043.30	16108.96	1680.02	1744.36

Table 4: Results of Duration Regressions for Early Retirement, using Household Income per Adult

	1	2	3	4	5	6
Hazard Ratios	Weibull	Weibull	Cox	Cox	Weibull	Weibull
					robust standard errors	
Self-rated health	**1.36**		**1.40**		**1.36**	
z	9.39		10.12		4.24	
P>|z|	0.000		0.000		0.000	
Health gradient		**0.70**		**0.66**		**0.70**
z		-6.39		-7.60		-3.95
P>|z|		0.000		0.000		0.000
Household income	**0.90**		**0.84**		**0.90**	
z	-1.58		-2.60		-1.32	
P>|z|	0.113		0.009		0.186	
Wealth	**1.02**	**1.02**	**1.00**	**1.01**	**1.02**	**1.02**
z	2.26	2.29	0.59	0.61	1.65	1.63
P>|z|	0.024	0.022	0.555	0.545	0.099	0.103
No employment	**1.64**	**1.40**	**1.38**	**1.14**	**1.64**	**1.40**
z	1.54	1.06	1.00	0.40	2.53	1.95
P>|z|	0.123	0.291	0.317	0.691	0.11	0.051
Public sector	**1.37**	**1.39**	**1.46**	**1.47**	**1.37**	**1.39**
z	5.04	5.18	5.97	6.09	4.08	5.20
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
Years of education	**0.97**	**0.98**	**0.96**	**0.97**	**0.97**	**0.98**
z	-2.47	-1.72	-3.71	-2.99	-1.70	-1.43
P>|z|	0.014	0.085	0.000	0.003	0.088	0.154
Female	**1.15**	**1.16**	**1.20**	**1.21**	**1.15**	**1.16**
z	2.32	2.51	3.00	3.11	8.02	4.81
P>|z|	0.021	0.012	0.003	0.002	0.000	0.000
Eastern residence	**1.15**	**1.10**	**1.24**	**1.20**	**1.15**	**1.10**
z	1.98	1.33	3.05	2.56	1.25	0.82
P>|z|	0.048	0.183	0.002	0.010	0.212	0.414
Unmarried	**1.16**	**1.18**	**1.14**	**1.15**	**1.16**	**1.18**
z	1.87	2.17	1.71	1.80	1.22	1.57
P>|z|	0.061	0.030	0.086	0.071	0.224	0.116
Year of birth	**0.89**	**0.89**	**0.94**	**0.94**	**0.89**	**0.89**
z	-14.53	-14.36	-7.91	-7.73	-7.98	-8.00
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
/ln_p	**1.05**	**1.09**			**1.05**	**1.09**
z	21.14	22.65			3.51	3.80
P>|z|	0.000	0.000			0.000	0.000
p	2.86	2.98			2.86	2.98
1/p	0.35	0.34			0.35	0.34
No. of subjects	10728	10728	10728	10728	10728	10728
No. of failures	1236	1236	1236	1236	1236	1236
Time at risk	16778048	16778048	16778048	16778048	16778048	16778048
Log likelihood	-827.59	-852.82	-8007.95	-8034.28	-827.59	-852.82
No. of obs	47255	47255	47255	47255	47255	47255
LR chi^2	390.48 (10)	340.04 (9)	297.30 (10)	244.63 (9)		
Prob > chi^2	0.000	0.000	0.000	0.000		
Akaike information cr.	1679.18	1727.64	16039.90	16090.56	1679.18	1727.64

Table 5: Results of Duration Regressions for Early Retirement, using Workers' Quintile Position on the Health Gradient

	1	2	3	4	5
Coefficients	Weibull	Weibull	Weibull	Weibull	Weibull
	——— Household income		———	— per capita	— per adult
Self-rated health / Household income	**1.34 / 0.84**				
z	8.20 / -2.25				
P>|z|	0.000 / 0.024				
Health gradient		**0.71**			
z		-4.87			
P>|z|		0.000			
Wealth	**1.02**	**1.02**	**1.02**	**1.01**	**1.01**
z	2.56	2.76	2.22	1.93	1.94
P>|z|	0.010	0.006	0.027	0.053	0.052
No employment	**1.48**	**1.30**	**1.54**	**1.61**	**1.55**
z	1.21	0.81	1.35	1.49	1.37
P>|z|	0.226	0.420	0.177	0.135	0.172
Public sector	**1.38**	**1.34**	**1.41**	**1.40**	**1.40**
z	5.07	5.21	5.51	5.38	5.38
P>|z|	0.000	0.000	0.000	0.000	0.000
Years of education	**0.98**	**0.98**	**0.96**	**0.96**	**0.96**
z	-2.13	-1.78	-3.49	-3.72	-3.53
P>|z|	0.033	0.074	0.000	0.000	0.000
Female	**1.13**	**1.13**	**1.15**	**1.18**	**1.16**
z	2.01	2.00	2.26	2.68	2.47
P>|z|	0.045	0.046	0.024	0.007	0.013
Eastern residence	**1.12**	**1.08**	**1.13**	**1.17**	**1.16**
z	1.63	1.02	1.80	2.27	2.09
P>|z|	0.103	0.308	0.072	0.023	0.037
Unmarried	**1.01**	**0.91**	**0.98**	**1.12**	**1.12**
z	0.08	-1.09	-0.26	1.50	1.44
P>|z|	0.937	0.275	0.795	0.132	0.149
Year of birth	**0.89**	**0.89**	**0.89**	**0.89**	**0.89**
z	-14.61	-14.70	-14.89	-14.73	-14.62
P>|z|	0.000	0.000	0.000	0.000	0.000
2nd quintile	**0.90**	**0.89**	**0.76**	**0.68**	**0.60**
z	-0.89	-0.97	-2.41	-3.25	-4.37
P>|z|	0.375	0.333	0.016	0.001	0.000
3rd quintile	**0.73**	**0.71**	**0.55**	**0.57**	**0.48**
z	-2.40	-2.65	-4.93	-4.63	-6.25
P>|z|	0.016	0.008	0.000	0.000	0.000
4th quintile	**0.82**	**0.78**	**0.57**	**0.62**	**0.53**
z	-1.46	-1.83	-4.63	-4.01	-5.50
P>|z|	0.145	0.067	0.000	0.000	0.000
5th quintile	**0.86**	**0.81**	**0.56**	**0.61**	**0.53**
z	-1.25	-1.68	-5.91	-5.11	-6.91
P>|z|	0.211	0.093	0.000	0.000	0.000
/ln_p	**1.04**	**1.08**	**1.05**	**1.06**	**1.07**
z	20.89	22.12	21.51	21.81	21.91
P>|z|	0.000	0.000	0.000	0.000	0.000
p	**2.84**	**2.93**	**2.87**	**2.89**	**2.90**
Log likelihood	-819.32	-841.09	-852.87	-859.46	-848.48
LR chi^2	407.04 (14)	363.50 (13)	339.99 (12)	326.81 (12)	348.77 (12)
Prob > chi^2	0.000	0.000	0.000	0.000	0.000
Akaike information criterion	1670.64	1712.18	1733.74	1746.92	1724.96

No. of subjects: 10728 No. of failures: 1236 Time at risk: 16778048 No. of obs. 47255

Table 6: Results of Shared Frailty Models for Early Retirement

	1	2	3	4	5	6
Hazard Ratios	Weibull	Weibull	Weibull	Weibull	Weibull	Weibull
	Household income		— per capita		— per adult	
Self-rated health	**1.36**		**1.37**		**1.34**	
z	9.30		9.21		8.20	
P>|z|	0.000		0.000		0.000	
Health gradient		**0.66**		**0.78**		**0.75**
z		-7.57		-3.70		-4.01
P>|z|		0.000		0.000		0.000
Household income	**0.80**		**0.92**		**0.95**	
z	-3.53		-1.25		-0.72	
P>|z|	0.000		0.211		0.472	
Wealth	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**	**1.02**
z	2.62	2.83	2.25	2.20	2.22	2.24
P>|z|	0.009	0.005	0.025	0.028	0.027	0.025
No employment	**1.50**	**1.32**	**1.64**	**1.49**	**1.60**	**1.37**
z	1.27	0.85	1.54	1.25	1.46	0.99
P>|z|	0.205	0.395	0.124	0.213	0.144	0.324
Public sector	**1.37**	**1.38**	**1.37**	**1.39**	**1.37**	**1.39**
z	5.01	5.11	5.03	5.21	5.03	5.21
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
Years of education	**0.98**	**0.98**	**0.97**	**0.97**	**0.97**	**0.98**
z	-1.98	-1.59	-2.68	-2.39	-2.70	-2.00
P>|z|	0.047	0.112	0.007	0.017	0.007	0.045
Female	**1.13**	**1.13**	**1.15**	**1.18**	**1.15**	**1.16**
z	2.07	2.08	2.34	2.73	2.25	2.42
P>|z|	0.039	0.038	0.019	0.006	0.025	0.015
Eastern residence	**1.11**	**1.07**	**1.16**	**1.13**	**1.17**	**1.11**
z	1.51	0.89	2.16	1.75	2.17	1.53
P>|z|	0.130	0.371	0.030	0.081	0.030	0.125
Unmarried	**1.03**	**0.94**	**1.15**	**1.17**	**1.14**	**1.16**
z	0.31	-0.81	1.83	2.03	1.72	1.95
P>|z|	0.758	0.417	0.068	0.042	0.086	0.051
Year of birth	**0.89**	**0.89**	**0.89**	**0.89**	**0.89**	**0.89**
z	-14.58	-14.71	-20.30	-20.21	-21.25	-20.97
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
/ln_p	**1.04**	**1.08**	**1.06**	**1.08**	**1.05**	**1.09**
z	20.77	22.27	30.17	31.70	31.76	33.97
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
/ln_theta	**-21.74**	**-15.68**	**-5.51**	**-5.51**	**-3.90**	**-3.80**
z	-0.03	-0.06	-3.25	-2.33	-4.29	-4.19
P>|z|	0.972	0.950	0.001	0.020	0.000	0.000
p	2.83	2.94	2.87	2.95	2.87	2.97
1/p	0.35	0.34	0.35	0.34	0.35	0.34
theta	3.61e-10	1.55e-07	0.004	0.004	*0.02	*0.02
LR of theta = 0: Chi^2	0.00 (1)	0.00 (1)	0.67 (1)	0.20 (1)	6.18 (1)	5.95 (1)
p-value	1.000	1.000	0.206	0.327	0.006	0.007
Log likelihood	-822.68	-845.08	-827.67	-861.08	-824.50	-849.84
Obs. per group: min	3657	3657	3357	3357	3365	3365
avg	9451	9451	9451	9451	9451	9451
max	27759	27759	28152	28152	28066	28066
LR chi^2	354.01 (10)	309.22 (9)	355.55 (10)	288.74 (9)	334.81 (10)	284.14 (9)
Prob > chi^2	0.000	0.000	0.000	0.000	0.000	0.000
Akaike information crit.	1669.36	1712.16	1679.34	1744.16	1673.00	1721.68

No. of subjects: 10728 No. of failures: 1236 Time at risk: 16778048 No. of obs. 47255 No. of groups: 5

Table 7a: Results of Stratified Parametric Duration Models for Early Retirement

Coefficients	1	2	3	4	5	6
	Weibull	Weibull	Weibull	Weibull	Weibull	Weibull
	Household income		— per capita		— per adult	
Self-rated health		**0.28**		**0.31**		**0.28**
z		7.90		8.63		7.79
P>|z|		0.000		0.000		0.000
Health gradient	**-0.36**		**-0.22**		**-0.29**	
z	-5.14		-3.16		-4.11	
P>|z|	0.000		0.002		0.000	
Household income		**-0.22**		**-0.08**		**-0.09**
z		-2.82		-1.06		-1.08
P>|z|		0.005		0.289		0.282
Wealth	**0.02**	**0.02**	**0.02**	**0.02**	**0.01**	**0.01**
z	2.52	2.35	2.00	2.01	1.94	1.88
P>|z|	0.012	0.019	0.045	0.044	0.052	0.060
No employment	**0.20**	**0.23**	**0.27**	**0.35**	**0.17**	**0.30**
z	0.61	0.93	0.85	1.09	0.52	0.92
P>|z|	0.54	0.355	0.395	0.276	0.603	0.358
Public sector	**0.33**	**0.32**	**0.33**	**0.32**	**0.33**	**0.33**
z	5.18	5.10	5.25	5.07	5.28	5.16
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
Years of education	**-0.02**	**-0.03**	**-0.03**	**-0.04**	**-0.03**	**-0.04**
z	-1.87	-2.20	-2.62	-2.92	-2.25	-2.93
P>|z|	0.061	0.028	0.009	0.003	0.024	0.003
Female	**0.15**	**0.15**	**0.19**	**0.17**	**0.17**	**0.16**
z	2.51	2.45	3.19	2.73	2.79	2.57
P>|z|	0.012	0.014	0.001	0.006	0.005	0.010
Eastern residence	**0.07**	**0.11**	**0.13**	**0.16**	**0.10**	**0.14**
z	0.92	1.47	1.86	2.27	1.41	1.99
P>|z|	0.355	0.141	0.063	0.024	0.159	0.046
Unmarried	**-0.10**	**-0.01**	**0.16**	**0.15**	**0.15**	**0.13**
z	-1.20	-0.14	2.07	1.90	1.96	1.73
P>|z|	0.232	0.887	0.038	0.057	0.050	0.083
Year of birth	**-0.11**	**-0.11**	**-0.11**	**-0.11**	**-0.11**	**-0.11**
z	-14.01	-13.86	-14.18	-14.19	-13.65	-13.77
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
2nd quintile	**-9.40**	**-9.17**	**-7.01**	**-6.90**	**-9.11**	**-8.91**
z	-3.86	-3.78	-2.78	-2.73	-3.77	-3.68
P>|z|	0.000	0.000	0.005	0.006	0.000	0.000
3rd quintile	**-14.50**	**-14.03**	**-10.92**	**-10.67**	**-12.47**	**-11.88**
z	-4.89	-4.78	-3.83	-3.75	-4.51	-4.31
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
4th quintile	**-21.22**	**-20.81**	**-17.42**	**-17.42**	**-27.81**	**-27.53**
z	-6.25	-6.13	-5.47	-5.46	-7.64	-7.58
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
5th quintile	**-23.43**	**-22.73**	**-21.52**	**-21.23**	**-24.59**	**-23.90**
z	-11.47	-11.18	-10.30	-10.19	-12.30	-11.98
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
Cons	**198.91**	**199.04**	**201.35**	**202.63**	**194.18**	**197.03**
z	12.55	12.58	12.62	12.72	12.23	12.45
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000

/ln_p						
2nd quintile	**0.53**	**0.54**	**0.39**	**0.40**	**0.51**	**0.52**
z	3.92	3.85	2.73	2.69	3.80	3.71
P>|z|	0.000	0.000	0.006	0.007	0.000	0.000
3rd quintile	**0.73**	**0.73**	**0.55**	**0.56**	**0.65**	**0.64**
z	5.21	5.09	3.88	3.81	4.70	4.50
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
4th quintile	**0.95**	**0.96**	**0.79**	**0.82**	**1.13**	**1.15**
z	6.94	6.84	5.73	5.74	8.68	8.59
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
5th quintile	**1.01**	**1.02**	**0.91**	**0.93**	**1.05**	**1.05**
z	9.15	8.96	8.11	7.99	9.81	9.56
P>|z|	0.000	0.000	0.000	0.000	0.000	0.000
Cons	**0.38**	**0.34**	**0.45**	**0.41**	**0.37**	**0.33**
z	3.36	2.98	3.94	3.45	3.42	2.99
P>|z|	0.001	0.003	0.000	0.001	0.001	0.003
No. of subjects	10728	10728	10728	10728	10728	10728
No. of failures	1236	1236	1236	1236	1236	1236
Time at risk	16778048	16778048	16778048	16778048	16778048	16778048
Log likelihood	-777.69	-777.69	-804.93	-772.85	-764.29	-742.52
No. of obs	47255	47255	47255	47255	47255	47255
LR chi^2	432.68 (13)	470.40 (14)	388.20 (13)	452.36 (14)	443.31 (13)	486.84 (14)
Prob > chi^2	0.000	0.000	0.000	0.000	0.000	0.000
AIC	1585.38	1587.38	1639.86	1577.7	1558.58	1517.04

Table 7b: Results of Stratified Semi-parametric Duration Model for Early Retirement

Coefficients	1 Cox Household income	2 Cox — per capita	3 Cox — per adult
Health gradient	**-0.40**	**-0.32**	**-0.38**
z	-5.88	-4.50	-5.25
P>|z|	0.000	0.000	0.000
Wealth	**0.01**	**0.00**	**0.00**
z	1.11	0.54	0.58
P>|z|	0.268	0.589	0.562
No employment	**0.07**	**-0.10**	**-1.16**
z	0.23	-0.29	-0.47
P>|z|	0.822	0.775	0.639
Public sector	**0.38**	**0.38**	**0.38**
z	5.96	5.97	5.88
P>|z|	0.000	0.000	0.000
Years of education	**-0.04**	**-0.05**	**-0.04**
z	-3.34	-3.83	-3.33
P>|z|	0.001	0.000	0.001
Female	**0.19**	**0.25**	**0.22**
z	3.11	3.98	3.47
P>|z|	0.002	0.000	0.001
Eastern residence	**0.15**	**0.21**	**0.17**
z	2.11	2.90	2.37
P>|z|	0.035	0.004	0.018
Unmarried	**-0.15**	**0.13**	**0.11**
z	-1.85	1.60	1.46
P>|z|	0.065	0.109	0.144
Year of birth	**-0.07**	**-0.07**	**-0.06**
z	-7.99	-7.93	-7.70
P>|z|	0.000	0.000	-0.000
No. of subjects	10728	10728	10728
No. of failures	1236	1236	1236
Time at risk	16778048	16778048	16778048
Log likelihood	-6349.31	-6402.27	-6377.50
No. of obs	47255	47255	47255
LR chi^2	194.52 (9)	188.77 (9)	185.00 (9)
Prob > chi^2	0.000	0.000	0.000
Akaike information criterion	12720.62	12826.54	12777.00

Table 8: Stratified Duration Regressions for Early Retirement in Accelerated Time Metric

Coefficients	1 Weibull	2 Weibull	3 Weibull	4 Weibull	5 Weibull
	Household income		— per capita	— per adult	
Self-rated health	**-0.08**				**-0.08**
z	-6.69				-6.67
P>|z|	0.000				0.000
Health gradient		**0.09**	**0.06**	**0.07**	
z		4.5	2.62	3.65	
P>|z|		0.000	0.009	0.000	
Household income	**0.05**				**0.11**
z	2.19				0.48
P>|z|	0.028				0.631
Wealth	**-0.01**	**-0.01**	**-0.01**	**-0.01**	**-0.01**
z	-2.39	-2.56	-2.19	-2.17	-2.05
P>|z|	0.017	0.010	0.029	0.030	0.040
No employment	**-0.11**	**-0.09**	**-0.11**	**-0.07**	**-0.09**
z	-0.93	-0.82	-0.97	-0.61	-0.84
P>|z|	0.352	0.410	0.331	0.544	0.402
Public sector	**-0.10**	**-0.96**	**-0.10**	**-0.10**	**-0.10**
z	-5.10	-5.07	-5.21	-5.27	-5.28
P>|z|	0.000	0.000	0.000	0.000	0.000
Years of education	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**
z	2.28	1.87	2.82	2.32	3.17
P>|z|	0.022	0.062	0.005	0.020	0.002
Female	**-0.04**	**-0.04**	**0.06**	**-0.05**	**-0.05**
z	-2.17	-2.31	-3.00	-2.85	-2.51
P>|z|	0.030	0.021	0.003	0.004	0.012
Eastern residence	**-0.03**	**-0.02**	**-0.03**	**-0.02**	**-0.03**
z	-1.22	-0.72	-1.49	-0.94	-1.56
P>|z|	0.222	0.472	0.136	0.346	0.118
Unmarried	**0.01**	**0.03**	**-0.04**	**-0.03**	**-0.03**
z	0.30	1.20	-1.69	-1.54	-1.21
P>|z|	0.765	0.230	0.090	0.124	0.225
Year of birth	**0.03**	**0.03**	**0.03**	**0.03**	**0.03**
z	8.72	8.81	8.94	8.67	8.68
P>|z|	0.000	0.000	0.000	0.000	0.000
2nd quintile	**0.11**	**0.11**	**0.17**	**0.22**	**0.23**
z	2.19	2.17	3.09	4.08	4.22
P>|z|	0.029	0.030	0.002	0.000	0.000
3rd quintile	**0.20**	**0.20**	**0.22**	**0.28**	**0.29**
z	3.76	3.78	3.95	5.19	5.26
P>|z|	0.000	0.000	0.000	0.000	0.000
4th quintile	**0.18**	**0.18**	**0.19**	**0.25**	**0.25**
z	3.30	3.38	3.52	4.76	4.78
P>|z|	0.001	0.001	0.000	0.000	0.000
5th quintile	**0.16**	**0.17**	**0.19**	**0.23**	**0.23**
z	3.25	3.30	3.63	4.59	4.62
P>|z|	0.001	0.001	0.000	0.000	0.000
Cons	**-53.98**	**-55.56**	**-54.99**	**-47.18**	**-50.32**
z	-7.52	-7.51	-7.69	-7.29	-7.40
P>|z|	0.000	0.000	0.000	0.000	0.000

/ln_p					
2nd quintile	**0.30**	**0.32**	**0.27**	**0.35**	**0.34**
z	3.61	3.82	3.12	4.07	4.02
P>|z|	0.000	0.000	0.002	0.000	0.000
3rd quintile	**0.40**	**0.42**	**0.32**	**0.42**	**0.39**
z	4.50	4.76	3.59	4.73	4.40
P>|z|	0.000	0.000	0.000	0.000	0.000
4th quintile	**0.55**	**0.58**	**0.49**	**0.71**	**0.68**
z	6.23	6.58	5.55	8.10	7.90
P>|z|	0.000	0.000	0.000	0.000	0.000
5th quintile	**0.58**	**0.62**	**0.56**	**0.67**	**0.64**
z	8.47	8.94	8.07	9.88	9.39
P>|z|	0.000	0.000	0.000	0.000	0.000
Cons	**0.71**	**0.72**	**0.74**	**0.70**	**0.69**
z	10.66	10.73	10.62	10.65	10.50
P>|z|	0.000	0.000	0.000	0.000	0.000
No. of subjects	10728	10728	10728	10728	10728
No. of failures	1236	1236	1236	1236	1236
Time at risk	16778048	16778048	16778048	16778048	16778048
Log likelihood	-772.59	-788.50	-812.36	-775.00	-756.62
No. of obs	47255	47255	47255	47255	47255
LR chi^2	322.10 (14)	290.27 (13)	284.93 (13)	281.82 (13)	318.59 (14)
Prob > chi^2	0.000	0.000	0.000	0.000	0.000
Akaike information crit.	1577.18	1607.00	1654.72	1580.00	1545.24

Table 9: Evolution of the Health Gradient over the Life-cycle for Alternative Measures of Income

Age groups	Household income		Per capita household income		Household income per adult		No. of obs.
	cons.	coef.	cons.	coef.	cons.	coef.	
21-25	3.10	-0.10	2.81	-0.82	2.86	-0.09	16448
26-30	2.92	-0.08	3.03	-0.10	2.98	-0.09	18936
31-35	3.76	-0.15	3.37	-0.12	3.96	-0.18	22144
36-40	4.35	-0.19	3.77	-0.15	4.68	-0.24	22440
41-45	5.24	-0.26	4.16	-0.18	5.08	-0.27	20851
46-50	5.60	-0.28	4.69	-0.22	5.30	-0.28	18390
51-55	**5.70**	**-0.28**	5.21	-0.26	5.50	-0.28	16769
56-60	5.67	-0.27	5.61	-0.28	**5.72**	**-0.30**	16306
61-65	5.24	-0.23	5.68	-0.29	5.69	-0.29	15667
66-70	5.04	-0.20	**5.82**	**-0.30**	5.82	-0.29	11439
71-75	4.68	-0.15	5.28	-0.22	5.26	-0.22	8467
76-80	4.41	-0.11	5.13	-0.19	5.10	-0.19	5451
81-85	3.99	-0.05	4.80	-0.14	4.79	-0.14	2809
86-90	3.11	0.05	4.81	-0.13	4.92	-0.14	1276
91-95	2.24	0.15	4.33	-0.07	4.41	-0.08	341
> 95	1.21	0.27	1.54	0.25	1.58	0.24	51

II. Figures

Figure 1: The Greater Decline of Health among Less Educated Workers

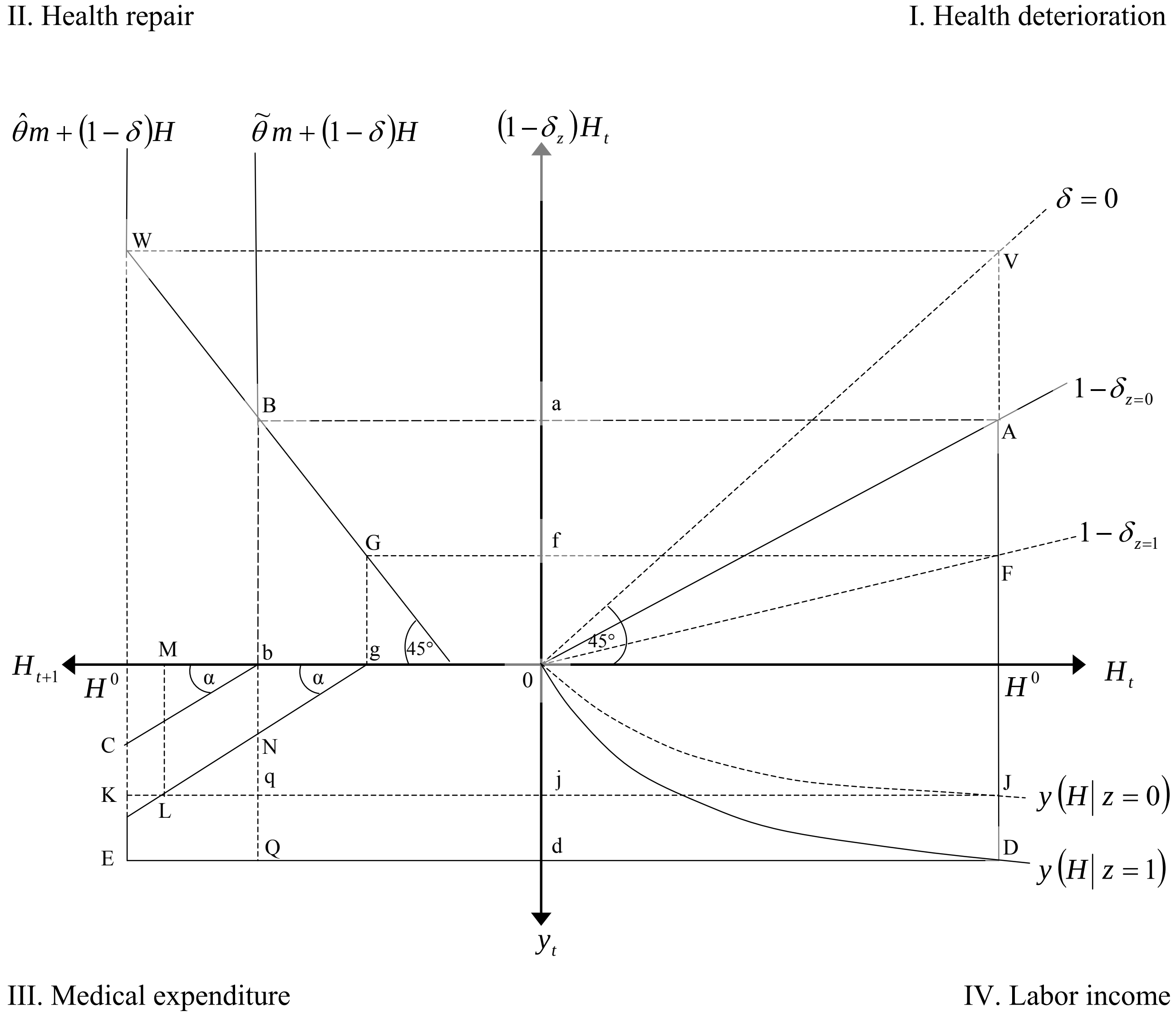


Figure 2: The Marginal Efficiency of Investments in Health and the User Costs of Health Capital

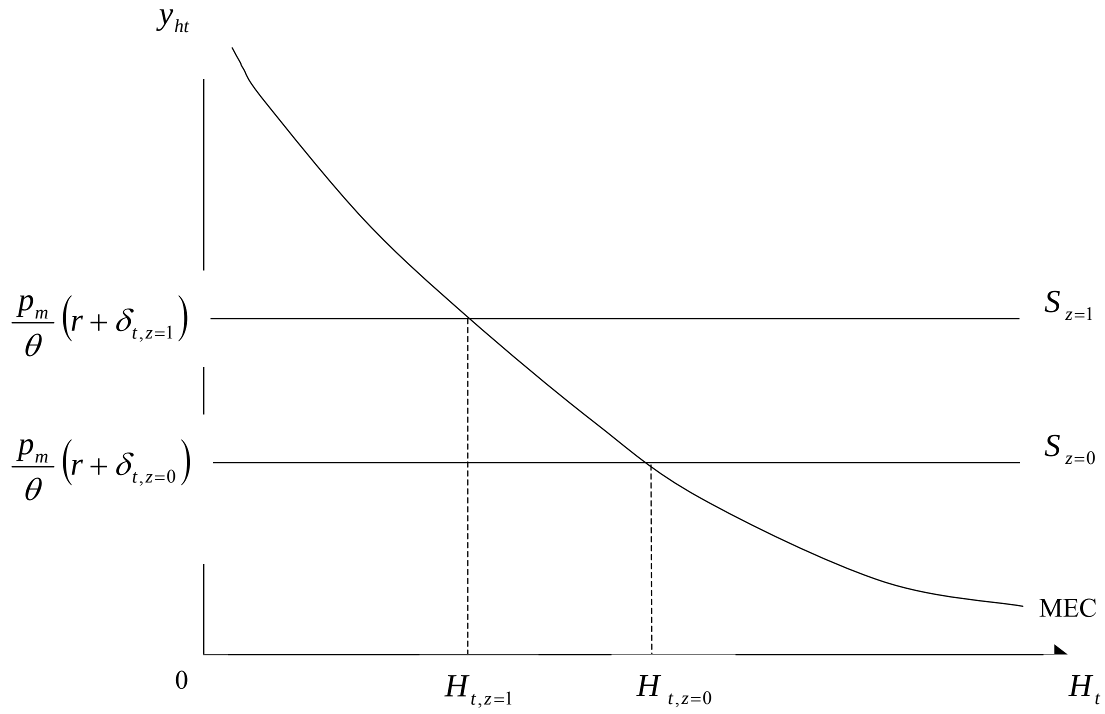
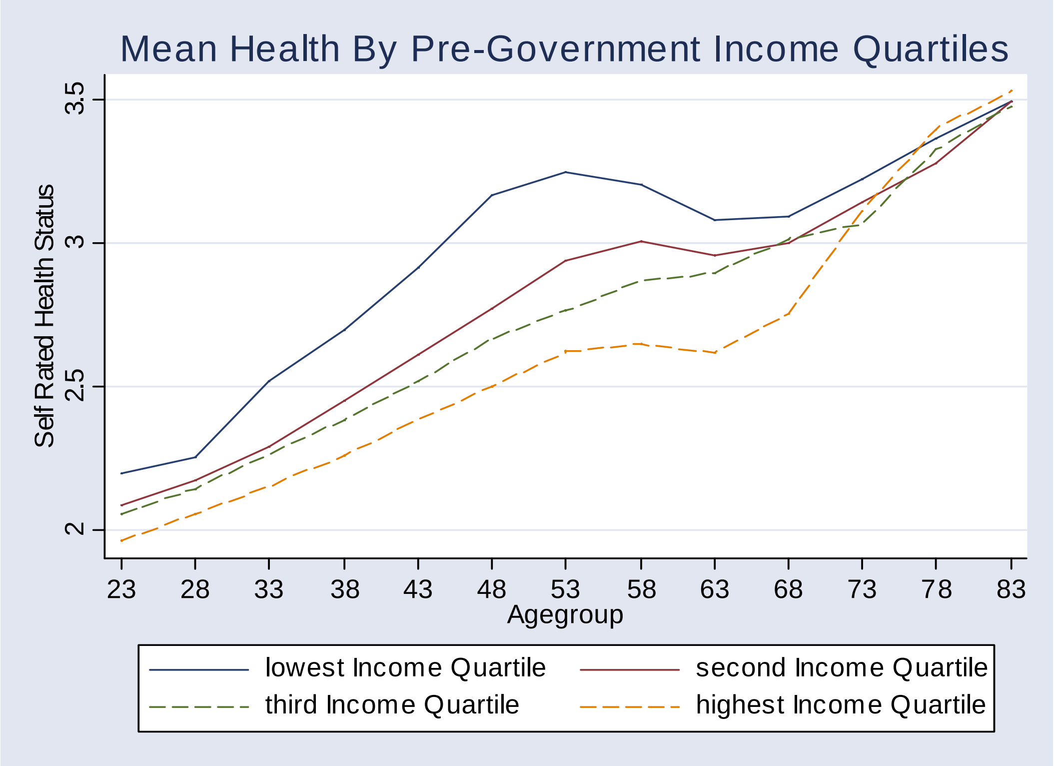


Figure 3: Agegroup-specific Self-rated Health Status over the Life-cycle

Panel A: Pre-government Household Income



Panel B: Post-government Household Income

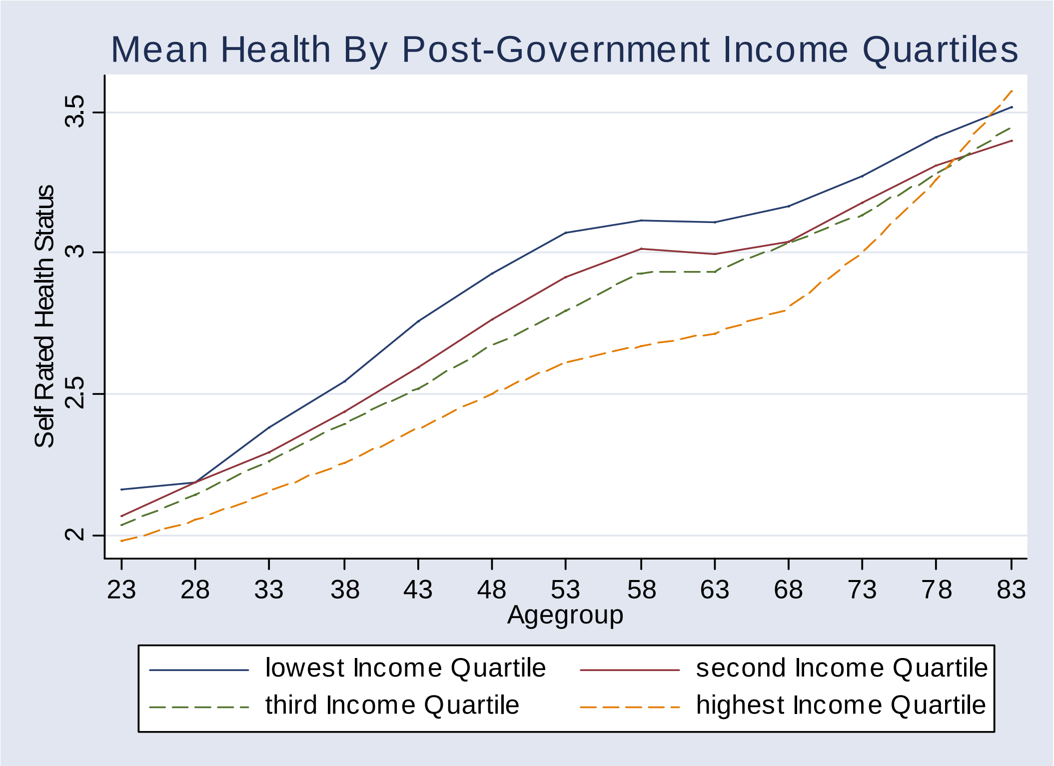


Figure 4: Evolution of the Health Gradient over the Life-cycle

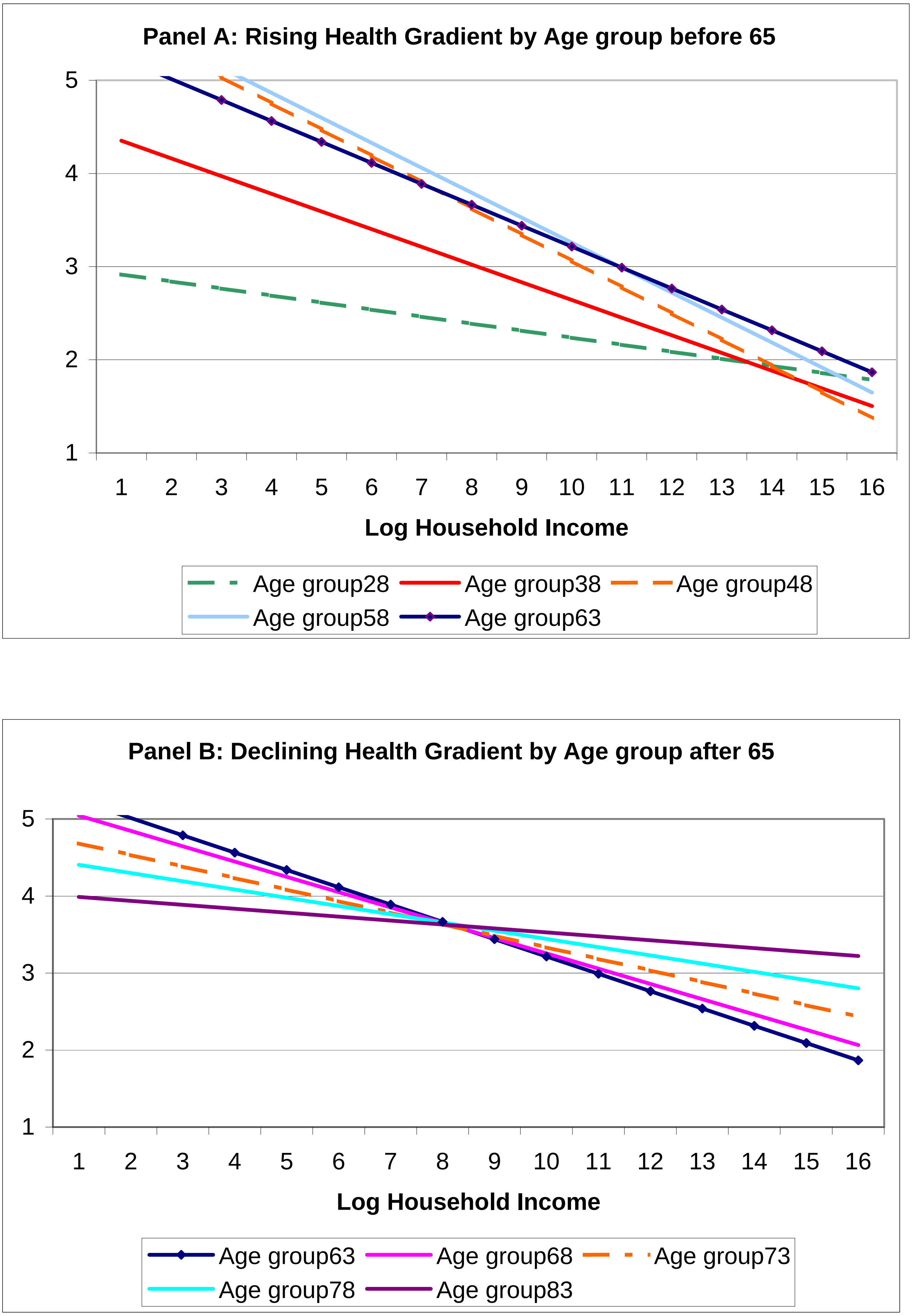
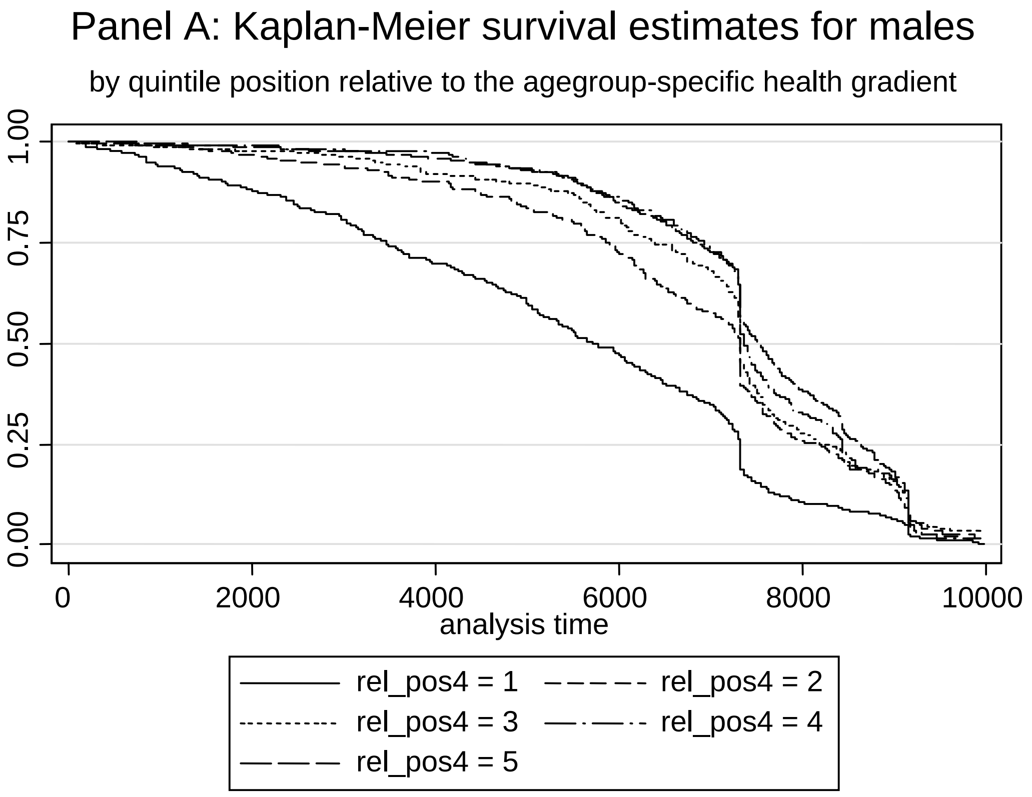


Figure 5: Non-parametric Duration Analysis of Retirement Timing in Germany



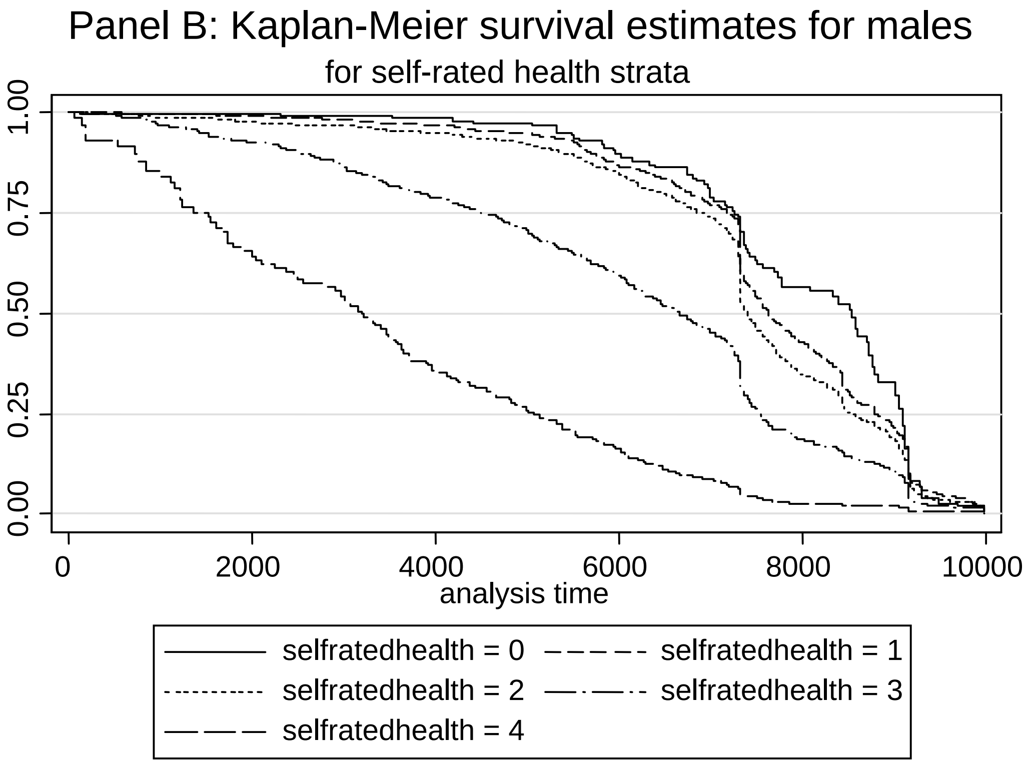


Figure 6: Predicted Hazard Rates of Early Retirement for Quintile Positions on the Health Gradient