OPTIMAL HEALTH AND RETIREMENT POLICIES AMID POPULATION AGING

Gisela Hostenkamp
Michael Stolpe*

Revised version of Kiel Working Paper 1428
http://www.ifw-kiel.de/pub/kap

July 2008

Abstract:

This paper develops a simple analytical framework in which optimal health and retirement policies amid population aging can be discussed. To be efficient, these policies must recognize and exploit the dynamic complementarities between the timing of retirement, the size of lifecycle labour income and pension payments and investments in health that individuals make, for example, by purchasing medical care and that society makes by advancing medical technology. We aim to show how the traditionally separate areas of health and retirement policy can be coordinated to achieve dynamic efficiency. Under fairly general assumptions, postponing the age of retirement and greater health spending are shown to be complements in the maximization of lifecycle utility. Mandatory retirement and pension policies can be used to induce voluntary health investments by individuals and improve society's incentives to adopt new medical technology. Leaving a hitherto optimal mandatory retirement age unchanged as new medical technologies improve the efficacy of healthcare would be inefficient. The aggregate willingness and ability to pay for medical care and technology will be greater, the higher an economy's per-capita income, suggesting large welfare gains from postponing the average age of retirement if investments in new medical technology target the quality of life and raise the productivity of people working past a long-established mandatory retirement age.

Keywords: Medical technology, Longevity, Health policy, Retirement age

JEL classification: I12, I18, J26

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

Gisela Hostenkamp
University of Southern Denmark
J.B. Winsløws Vej 9, DK–5000 Odense C
Telephone: +45-6550-3886
Email: gih@cast.sdu.dk

Acknowledgement: This paper is part of the Kiel Institute's research project on "The Role of Medical Technology in the Demographic Transition of the 21st Century," funded by Pfizer, New York.

I. INTRODUCTION

The prospect of rapid population aging has made pay-as-you-go pension systems in many countries look fiscally unsustainable, unless they are fundamentally reformed. Similar claims are often made for healthcare although the demographic impact on per-capita health spending is less clear. Investments in health and new medical technologies may in fact help to compress the time that people typically spend in morbidity before they die. Moreover, many workers may be able to stay in their jobs well beyond the current retirement age and remain contributors instead of becoming recipients of the pension system. These opportunities are often ignored and new medical technology is widely seen as an inefficient contributor to rising per-capita health spending – not as a solution, but as an additional burden amid population aging. "Can we afford to live longer in better health?" is the title of a recent generational accounting study for the EU-15 countries by Westerhout and Pellikaan (2005) in which they suggest: "To mitigate the effect of ageing on healthcare expenditures, (…) healthcare budgets may be frozen for several years, or expenditures cut, so that healthcare expenditures grow at a slower rate than GDP for several years."

Our paper argues against such a pessimistic outlook and develops a simple optimization model to analyze the normative implications of new medical technologies, endogenous longevity and improvements in the quality of life for public pension and health policies. To be efficient, these policies must recognize and exploit the dynamic complementarity between the timing of retirement, the size of pension payments and advances in medical technology: The aggregate ability and willingness to pay for medical technology will be greater, the higher an economy's per-capita income, suggesting large welfare gains from postponing the average age of retirement if investments in new medical technology target the quality of life and raise the productivity of people working past the currently effective retirement age. A successful coordination of health and retirement policies will be essential to keep public-sector debt at sustainable levels during the 21st century, as a recent empirical study of pension promises in 30 OECD countries (Queisser and Whitehouse, 2006) suggests that further cuts in mandatory pensions would be politically unacceptable in many of these countries.

How much investment in health would be justified by new opportunities to expand people's working lives? And how should opportunities to improve health affect the rules about retirement age? We discuss these questions for a representative agent with a finite life, ignoring trades between generations, such as bequests, and detailed social security design issues that are often studied within overlapping generation models in the spirit of Samuelson (1958) and Diamond (1965). We also ignore the role of fertility and introduce population aging simply by extending individual lives.

The basic elements of our model can be summarized in terms of price and income effects on the timing of individual retirement, as illustrated by means of the utility indifference curves of a representative worker in Figure 1. In the simplest case, each worker has a fixed time budget over the lifecycle, such as T_0T_0, the only savings motive is consumption smoothing and there is no discounting. Assuming non-satiation in leisure and monetary income, an optimum is reached at *A*. Next, we assume retirement stops the health decline associated with working so that earlier retirement generates additional lifetime; Hostenkamp and Stolpe (2008) find that German workers retiring before 60 years of age even improve their health within three years after the event. This implies a higher relative price of work, illustrated by the new budget constraint T_0T_1. The new optimum, at *B*, may imply more time at work in absolute terms, but the substitution effect lets the lifetime share of work decline, depending on preferences and the relative size of income and substitution effects. In the graph, *B* lies to the left of the ray, not actually drawn, through the origin and *A*, illustrating the robust finding of many empirical studies that the substitution effect dominates (Dwyer and Mitchell, 1999, p. 175). Finally, other health investments, such as across-the-board improvements in medical technology, move the original budget constraint from T_0T_0 to T_2T_2, equivalent to a pure income effect that leads to a new optimum at *C*, where both leisure and working time are expanded in similar proportion. However, if an immovable mandatory retirement age had fixed working time at the level of point *A*, the welfare gains from better medical technology would obviously be smaller than those in *C*. This illustrates the basic argument for the coordination of health, pension and retirement policies amid population aging.

The issue is related to several strands of prior literature. A need for coordination is clearly implied by Philipson and Becker (1998) who model the influence of public pension and healthcare policies on the private demand for health and longevity. They argue that mortality-contingent claims, by making wealth dependent on the duration of life, can induce behavior and investments with a positive effect on longevity, at the expense of quality of life. This distortion is often compounded by public pension programs' high implicit rates of wage taxation that pull older workers out of the labor force regardless of productivity (Gruber and Wise, 1998). This may have been a small problem when pensions merely served as insurance against living longer than the average worker, but many of today's elderly have substantial unused productivity potential well beyond the normal statutory retirement age, at 65 in most European countries. Macroeconomic estimates, such as Becker et al. (2005) and Nordhaus (2003), suggest we would underestimate economic growth by half if we did not include gains in health and longevity in our measure of output. Many of these gains seem to have been privately appropriated by workers spending an ever larger part of their lives in retirement so that microeconomic estimates of the health effect on per-capita income, such as in Weil (2007), are likely to underestimate the social returns to better health substantially. Larger gains would have been realized had health and retirement policies already been coordinated as we propose in this paper.

The rest of our paper discusses the economic rationale for a mandatory retirement age (section 2), introduces a simple model of the influence of retirement timing on health investments (section 3), uses it to analyze the optimal coordination of health and retirement policies (section 4), and concludes with a general discussion in section 5.

II. RATIONALE

A valid economic rationale for a statutory, or even mandatory, retirement age and its coordination with health policy presupposes that retirement and lifecycle health are interrelated, that unregulated individual demand for retirement may be distorted and that uncoordinated private demand for healthcare may be insufficient. To define the underlying rationale for retirement per se, we distinguish between individual choice and legal mandates, including the regulation of retirement age. Public pension systems throughout Europe and elsewhere are usually associated with sufficient financial penalties for working past the statutory retirement age that we can take it as mandatory for the purpose of our analysis.

Consider the individual perspective first: In the absence of mandatory retirement and death, people would *have* to work forever, unless they inherit wealth or succeed in saving sufficiently to live from a perpetual annuity. Yet, mortality is real and people do not seem to want an equal distribution of work and leisure across all periods of their finite lifetime, although this would be consistent with perfect consumption smoothing even in the absence of capital markets, provided productivity is constant. In reality, productivity cannot be constant when finite lifetimes are associated with the depreciation of human capital, in the form of health, skills and acquired knowledge. The irreversible depreciation of health, often accelerated by effort at work, is an implication of imperfect medical technology. Skills and acquired knowledge become obsolete as the economy's general stock of knowledge grows. As an individual's remaining lifetime declines, so does the potential payoff from acquiring new skills and knowledge. In anticipation of these declining opportunity costs, people may rationally defer a large part of total lifetime leisure to the end of life.

To establish an additional rationale for *mandatory* retirement and for a general mandatory retirement *age* presupposes that the individual rationale is somehow incomplete or inefficient, for example due to social interactions. In this vein, Sala-i-Martin (1996) hypothesizes that public pensions are linked to retirement from the workforce because positive externalities in the average stock of human capital imply a negative effect of the elderly's lower-than-average skills on the productivity of younger workers. Retirement of older workers helps to raise aggregate output, unless their negative impact can be internalized in other ways, for example through a lower wage. We see a similar rationale for a mandatory retirement age in countries with equal access to universal healthcare, based on the *pecuniary* externalities generally associated with third-party payer systems: Without mandatory retirement, workers might overwork their bodies by staying in the workforce too long, anticipating to bear only the burden of irreparable health shocks, not the direct pecuniary cost of care.

Important *pecuniary* externalities are also associated with investments in new medical technology, which is often knowledge-intensive, provided under increasing returns to scale and has characteristics of a public good. The more people contribute in line with their true individual willingness-to-pay, the closer will society's aggregate ability to provide financial incentives for innovators come to match medical technology's true social payoff. A mandatory retirement age that is adjusted in anticipation of gains in health and longevity may often be the simplest and most flexible way of aggregating people's individual willingness-to-pay, essentially a reversal of in-kind transfers in lieu of a direct tax. Correctly anticipating future health and mortality is difficult as we do not have a suitable theoretical model of endogenous mortality that quantifies the influence of the many relevant factors, such as work, income, education, nutrition, sanitation and access to medical care, whose relative weight may vary with a person's age, cohort and place and time of living. Oeppen and Vaupel (2002) argue past predictions were often misled by the assumption that biological barriers impose an immovable ceiling on average longevity, whereas best-performance life expectancy across all countries with reliable population statistics has in fact shown a stable long-term trend adding a quarter of a year annually for more than 150 years. At least in developed countries, the effectiveness of medical technology appears to have replaced nutrition, sanitation and working conditions as the ultimate constraint on life expectancy (Cutler et al., 2006), with the bulk of recent life expectancy gains concentrated in the years after 65.

III. A SIMPLE MODEL OF RETIREMENT AND HEALTH

In this section, individuals are assumed to make health investments in response to non-anticipated changes in exogenous variables that determine the optimal level of health. Our focus is on the role of retirement age and we show how raising a given mandatory retirement age tends to increase workers' incentives to invest in their health. Pay-as-you-go and other defined-benefit pension systems typically make the eligibility to receive pension payments contingent on leaving the labor force at a statutory retirement age and impose various restrictions on early retirement, implying substantial deviations from actuarial fairness for those seeking early retirement or working past the statutory retirement date. In a stark simplification, we assume the presence of actuarially fair annuities, so that no capital market inefficiency complicates our argument, and rule out any deviation from the statutory retirement age, which we therefore denote as mandatory. Under these assumptions, increasing the age of mandatory retirement and greater health spending turn out to be complements in the consumer's utility maximization.

There are two ways of looking at this complementarity. On the one hand, a higher mandatory retirement age may make more lifetime income available to finance consumption in a given number of periods, in each of which marginal utility of consumption declines. Some of the additional income will hence be used for health investments and for consumption in the additional years of life that this may generate. The incentives to make such investments would be lower if retirement were mandatory at an unchanged age as they would come at the expense of consumption in earlier years. On the other hand, better health counteracts the detrimental effect of later retirement on survivorship. Raising the mandatory retirement age without allowing workers to improve their health might lead to an excessive working life, in which the marginal benefit of continuing to work falls short of marginal costs. Investments in health address this imbalance by raising survivorship and expected lifetime income and by lowering the mortality benefit from retirement at any given age. As both the level of additional wage income from postponing retirement and its derivative with respect to retirement age, R, increase, the latent demand for retirement before the new mandatory age declines. We assume a higher R always raises per-period income since it generates additional lifetime income for a shorter lifespan, given that working longer increases mortality. Nonetheless, because the objective is to maximize lifetime utility which grows with the duration of life, a negative function of *R*, it is generally not optimal to raise *R* until life's end.

To keep the analytical framework as simple as possible, we adapt Hall and Jones' (2007) stylized model of rising health spending amid per-capita income growth. Although they do not address the retirement issue, they introduce a utility function in which average utility is always greater than marginal utility, thus motivating the pursuit of longevity that links retirement timing, *R*, and health investments, *h*. To this end, they add a constant flow of utility to the time-additively separable per-period utility from consumption. Preferences distinguish between the flow of per-period utility, u, a function of consumption *c*, and lifetime utility, the product of per-period utility and life expectancy *g*. As Rosen (1988) first saw, the level of u affects the optimal intertemporal allocation because adding a constant raises the value placed on longevity relative to the instantaneous consumption of goods and ensures an income elasticity of health demand above one (Dormont et al, 2007, pp. 60): rising income is associated with a falling elasticity of utility with respect to consumption, η_c, *relative* to the elasticity of health with respect to healthcare spending, η_h. Health spending may therefore be greater than the financial-wealth-maximizing level and the optimal retirement may take place before the financial-wealth-maximizing retirement age.

In the equilibrium of our model, the representative agent faces the same mortality hazard at all ages, $1/H$, defined as the inverse of life expectancy or health status, *H*, which is determined by the health production function $H = g(R,h)$ with $\partial g(\cdot)/\partial R < 0$, $\partial^2 g(\cdot)/\partial R^2 < 0$, $\partial g(\cdot)/\partial h > 0$, $\partial^2 g(\cdot)/\partial h^2 < 0$, and $\partial^2 g(\cdot)/\partial R \partial h > 0$; the latter indicates that greater health spending reduces the detrimental effect of later retirement on health. Although an age-invariant mortality hazard may not seem realistic, especially when *R* affects mortality at higher ages, we think of it as a limiting case in the presence of competing risks. As *R* is raised, we hypothesize there will be an accelerating increase in mortality during the additional working time. But as these changes are rationally anticipated in our model, younger workers make adjustments with regard to other hazardous activities that lead to a new equilibrium distribution of mortality risk throughout life, as predicted by the theory of competing risks (Dow et al., 1999): complementary effects between multiple causes of death, such as hazards before and after a given age, operate to equalize the occurrence of the causes. More specifically, when later retirement raises mortality in old age, workers no longer put so much effort into surviving to that age and let mortality rise at younger ages, too. For example, they may choose better paid jobs with a higher mortality risk.

In the simplest version of the model, Hall and Jones (2007) assume a constant flow of income per period, *y*. In our case, this would imply *y* is unaffected by health spending, *R* is fixed, and the consumer maximizes expected per-period utility $g(R,h)u(c)$ subject to the per-period budget constraint, $c+h=y$. The optimal allocation would then equate the ratio of health spending to consumption with the ratio of the elasticities of the health production function and the flow utility function so that the optimal health share, $s \equiv h/y$, satisfies $s^*/(1-s^*) = h^*/c^* = \eta_h/\eta_c$, where $\eta_h \equiv g'(h)h/H$ and $\eta_c \equiv u'(c)u/c$. We can think of this solution as a theoretical benchmark.

More realistically, health spending that lets mortality decline does have an impact on per-period income. To some extent, it will change even if the age of retirement is held constant because on the one hand active workers are less likely to die before reaching that age and on the other hand time spent in retirement tends to be longer as life expectancy rises. Taking these two opposite effects into account, the first-order conditions imply $h/c = (\eta_h/\eta_c)/[1-(\partial y(\cdot)/\partial g)(\partial g(\cdot)/\partial h)]$, which shows optimal health spending over consumption to be greater (or smaller), the larger (or smaller) the marginal effect of greater health spending on per-period income. The overall effect comprises the effect of changing health spending on health, $\partial g(\cdot)/\partial h$, and of changing health on income, $\partial y(\cdot)/\partial g$. While the former effect is assumed to be always positive, it is not a priori clear how lower mortality affects per-period income. With equal distribution of lifetime income across periods and constant productivity up to retirement, the net effect depends on the relative gains in expected lifetime before and after retirement. Formally, life expectancy is $\int_0^R e^{-t/H}dt = (1-e^{-R/H})H$ before and $\int_R^\infty e^{-t/H}dt = He^{-R/H}$ after retirement. Depending on the given age of retirement, the gain in expected time before retirement may be larger or smaller than the gain in expected time after retirement. More specifically, when *R* is high, the additional lifetime income due to declining mortality before retirement may be large enough to prevent a fall in per-period income even if a rising *h* prolongs overall lifetime. However, if the gains in life expectancy are large, a parallel postponement of the retirement age may be required to generate all the additional lifetime income that is needed to maintain a given per-period income flow. If *R* is increased in response to declining mortality, this may ensure $\partial y/\partial g > 0$ in addition to $\partial g(\cdot)/\partial h > 0$.

The optimal income share of health spending when it does impact on per-period income, is determined by $s = \eta_h / [\eta_c(1-(\partial y/\partial g)(\partial g/\partial h)) + \eta_h]$. For reasons of system's stability, this presupposes that the marginal effect of greater health spending on per-period income is smaller than one. In the benchmark case of no impact that Hall and Jones (2007) study, we have $(\partial y/\partial g)(\partial g/\partial h) = 0$, and the optimal health share collapses to $s^* = \eta_h/(\eta_c + \eta_h)$. A positive impact of health spending on per-period income implies a larger health share, whereas a negative impact would imply a smaller health share than the optimal share in the limiting case.

To focus on the implications of changes in retirement age, we now assume that total lifetime income is proportional to the expected value of the duration of working life, $(1 - e^{-R/H})H$. The individual budget constraint under the per-period flow of income, *y*, is now given by $c + h = y(R, H) = y(R, g(R,h)) = y\left[(1 - e^{-R/g(R,h)})g(R,h)\right]$, where the budget spent in a given period is a function of *H* and *R*. Normalizing gross wage income in each period before retirement to one, we can express the available net per-period income flow as $y = 1 - e^{-R/g(R,h)}$ since *g* is equal to life expectancy.

To motivate our assumption of equal per-period income across working life and time in retirement, we rely on the absence of time preference, the absence of any direct impact of retirement on utility and the presence of full and fair insurance against the risk of running out of resources when death comes later and of leaving unspent resources when death comes earlier than expected. For analytical convenience, we further assume that equal health spending across periods is optimal, say, because the effectiveness of health spending is age-invariant and because it is effective only in the period in which it is made. These assumptions imply some shifting of resources from the time of work to the time in retirement. We assume this takes place at the social level and do not explicitly model it here. By contrast, Hall and Jones (2007), whose model has no retirement, rule out any opportunity to shift resources between periods and thus bypass the need for actuarially fair annuities.

To determine the impact of increasing *R* on *h*, we set up the Lagrangian $L = g(R,h)u(c) + \lambda(1 - e^{-R/g(R,h)} - h - c)$ and derive the first-order condition for *h*

$\partial L/\partial h = (\partial g(\cdot)/\partial h)u(c) + g(\cdot)\partial u(c)/\partial c\left[-e^{-R/g(R,h)}(\partial g(\cdot)/\partial h)\, R/g(R,h)^2 - 1\right] = 0$, where the Lagrange multiplier takes the form $\lambda = g(R_0, h_0)\partial u(c_0)/\partial c$. Denoting $\partial u(c_0)/\partial c$ as u_c and applying the implicit function theorem, we find

$$\frac{dh}{dR} = -\frac{Ru_c\frac{\partial g}{\partial h}\left((R - g(\cdot))\frac{\partial g}{\partial R} - g(\cdot)\right) + g(\cdot)^2 u_c\left(\frac{\partial g}{\partial h} + R\frac{\partial^2 g}{\partial h \partial R}\right) + e^{-R/g(\cdot)}g(\cdot)^3\left(u_c\frac{\partial g}{\partial R} - u(c)\frac{\partial^2 g}{\partial h \partial R}\right)}{-Rg(\cdot)\frac{\partial^2 g}{\partial h^2}u_c\left(\frac{\partial g}{\partial h}\right)^2 + R^2 u_c\left(\frac{\partial g}{\partial h}\right)^2 + Rg(\cdot)^2 u_c\frac{\partial^2 g}{\partial h^2} + e^{-R/g(\cdot)}g(\cdot)^3\left(u_c\frac{\partial g}{\partial h} - u(c)\frac{\partial^2 g}{\partial h^2}\right)}.$$

The denominator is likely to have a positive sign since all its parts, with the exception of the third addend, are positive and the third addend is not likely to be larger in absolute value than the sum of the other three. In the numerator, the first and third parts have negative signs. Only its middle addend is positive, but while it may be larger in terms of absolute value than the first and third addend separately, it is not likely to be larger than the two other addends together. Given that numerator and denominator almost certainly have different signs, the overall effect of increasing *R* on *h* is likely to be positive.

Figure 2 illustrates the static model of Hall and Jones (2007) in the original version with only one control, namely health spending (curve *V*), and our adaptation with an additional control, the time of retirement (curve *V'*). When *R* is fixed, every exogenous increase in per-period income shifts the budget constraint away from the origin, such as from $T_R T_R$ to $T_2 T_2$. The model implies that people do not increase their spending on health and other consumption proportionally, but that the share of health in income increases, as indicated by moving from point *D* to *E* on the health spending expansion curve *V*, well above the 45°-line. By contrast, curve *V'* illustrates the implications for healthcare spending of an individual choice of retirement timing as an additional way to improve health. If the budget constraint were fixed, such as in $T_2 T_2$, the individual would use the retirement option to spend less on healthcare and more on other consumption, as indicated by moving from point *E* to *F*. The new allocation in *F* lies on the new health spending expansion curve V' indicating a less rapidly rising health share as income rises. In this case, increasing c seems to be desirable because the health gain from early retirement lowers the marginal product of additional healthcare spending. However, the lifetime budget is also lower so that only point *G* on the new budget constraint $T_R T_R$ can be reached. While health spending is clearly lower in *G* compared to *E*, the effect on other consumption spending is ambiguous.

Finally, Figure 2 may also be interpreted in terms of later retirement. Per-period income rises as retirement is postponed, making additional health spending that increases life expectancy more attractive than additional consumption whose marginal utility declines. Similarly, when health spending is added as an additional choice to increase longevity, the marginal gain from continued work at a given age rises and keeping the retirement age constant becomes more costly in terms of forgone opportunities, as indicated by moving from *G* to *E*. Later retirement and a rising health share in income are thus again shown to be complements.

IV. OPTIMAL COORDINATION

The role of greater income as a source of greater health spending may give rise to positive feedback and may thus facilitate sustainable population aging. If the government raises a mandatory age of retirement so that per-period income rises, we predict that health spending, too, will rise, even more than proportionally, and call for a further rise of the optimal retirement age. By the envelope theorem, we know that the impact of health spending on the new equilibrium per-period income will generally be larger if the retirement age is optimally adjusted in response to changes in health, regardless of whether the timing of retirement itself has an effect on health. If later retirement does increase mortality, this effect will be small for small changes in retirement timing as it concerns only the time after the original retirement age. Provided healthcare is sufficiently productive, there will be at least some scope to boost lifetime income by delaying retirement in response to greater health spending without fully cancelling the gains in health that this spending is intended to bring about. Investments in health, or exogenous shocks that improve the productivity of medical technology, can thus create a virtuous circle in which retirement age and incomes rise as mortality declines. This raises the issue of the optimal speed and coordination of the process, with normative implications for government policy.

To demonstrate the benefits of coordination, we compare the value of the indirect utility function for increases in h under an exogenously fixed R with its value when h is accompanied by an optimal increase in R. For given values of the set of exogenous parameters $\alpha = \{H_0, \eta_c, \eta_h, \eta_R\}$ and choice variables $x = \{R, h, c\}$, the indirect utility function $\phi(\alpha)$ is defined as the maximum value of $f(x,\alpha) = H\,u(c)$ subject to a set of constraints $\varphi = \left\{R - R_0 = 0,\ g(R_0, h, H_0) - H = 0,\ y = 1 - e^{-R/g(\cdot)} = c + h\right\}$. The greater potential utility when retirement is flexible is an example of the well-known Le Châtelier effects, implying that long-term demand for healthcare is more elastic than short-term demand. The short term is defined as a situation in which the mandatory retirement age is held constant at the age appropriate before one or more of the exogenous parameters are changed. The short-term indirect utility function, denoted ${}^{2}\phi(x,\alpha)$, thus features two relevant constraints – namely one on the generation of per-period income, using health spending as an input in the production of health, and the other on the mandatory retirement age, R_0. The long-term indirect utility function ${}^{1}\phi(x,\alpha)$ has only one constraint – on the generation of per-period income – and is therefore more concave around $R = R_0$ than ${}^{2}\phi(x,\alpha)$, implying different comparative statics around the original solution, at which the constraint on *R* is just binding in ${}^{2}\phi(x,\alpha)$. For the technical background, see Silberberg (1990, pp. 216–222).

To proceed, we need the unconstrained optimal choices for *c* and in particular for *R* and *h* as functions of the exogenous parameters. As before, the consumer maximizes $g(R,h)\,u(c)$ with respect to *c*, *h* and *R* and subject to $c + h = y\,(R,\, g(R,h)) = 1 - e^{-R/g(R,h)}$. While we do not impose a specific functional form for $g(R,h)$, we assume that *g* is monotonically *increasing* in *h* and monotonically *decreasing* in *R*, as explained in section 3, to ensure the maximization problem is well-defined. Health spending essentially run into diminishing marginal returns because medical technology is imperfect and cannot repair all health shocks fully. Otherwise, health spending might be raised until mortality drops to zero so that workers could work forever, without retiring. In this case, the maximand of our model would be unbounded, the problem ill-defined. Imperfect medical technology ensures mortality is always positive and retirement is optimal at some finite age, given that mortality rises monotonically and at an increasing rate with *R* if *h* is held constant.

Taking derivatives of the Lagrange function $L = g\,(R,h)u(c) + \lambda\left(1 - e^{-R/g(\cdot)} - c - h\right)$ yields the first-order conditions

$$\partial L/\partial c = g(\cdot)\partial\, u(c)/\partial c - \lambda = 0\,,$$

$$\partial L/\partial h = (\partial g(h)/\partial h)u(c) - \lambda\left(e^{-R/g(\cdot)}R(\partial g(\cdot)/\partial h)\big/ g(\cdot)^2 + 1\right) = 0\,,$$

$$\partial L/\partial R = (\partial g(\cdot)/\partial R)u(c) - \lambda\, e^{-R/g(\cdot)}\left(R\,(\partial g(\cdot)/\partial R)\big/ g(\cdot)^2 - 1/g(\cdot)\right) = 0\,,$$

and $1 - e^{-R/g(\cdot)} - c - h = 0$.

Using $\lambda_0 = g(R_0, h_0)\partial u(c_0)/\partial c$ and $e^{-R_0/g(\cdot)} = 1 - c_0 - h_0$, we can write the two first-order conditions for *h* and *R* as

$$\partial L/\partial h = u(c)\partial g(\cdot)/\partial h - g(R,h)\,\partial u(c)/\partial c\left[1 + (1-c-h)R\,(\partial g(\cdot)/\partial h)/g(R,h)^2\right] = 0$$

$$\partial L/\partial R = u(c)\partial g(\cdot)/\partial R - g(R,h)\,\partial u(c)/\partial c\,(1-c-h)\left[R\,\partial g(\cdot)/\partial R/g(R,h)^2 - 1/g(R,h)\right] = 0$$

Combining and using $u'(c) = \eta_c\, u/c$ where the elasticity of marginal utility, γ, is assumed constant, $\partial g(\cdot)/\partial h = \eta_h\, g/h$ where the elasticity of health with respect to health spending, η_h, is declining, and $\partial g(\cdot)/\partial R = \eta_R\, g/R$ where the negative elasticity of health with respect to retirement age, η_R, is increasing in terms of its absolute value, we have $(\eta_R\, g/R)/(\eta_h\, g/h + g\,\eta_c\, u/c) = (\eta_c\, u/c\,(\eta_R - 1)g^2)/(R\eta_h\, g/h)$, which can be further simplified to $\eta_R/(\eta_h/h + \eta_c\, u/c) = \eta_c\, u/c\,(\eta_R - 1)g/(\eta_h/h)$. It follows that per-period health spending,

$$h = -c\,\eta_h(c\,\eta_R - (\eta_R - 1)g\,u\,\eta_c)\,/\,(g\,u^2\,\eta_c^2(\eta_R - 1)),$$

can be represented as a function of three important elasticities, η_c, η_h, and η_R. The optimal value of *h* is larger the lower both the initial levels of health, *g*, and utility, *u*, the lower η_c, the larger – in terms of absolute value – η_R and the larger η_h. The ratio of *h* to *c* depends on the same set of exogenous parameters and rises with *c* itself: $\partial(h/c)/\partial c > 0$, as shown in Figure 2.

Deriving explicit choice functions for *R*, *h* and *c* in terms of the exogenous parameters only is difficult without further specifying the functional form of $g(R,h)$. Moreover, the first-order condition $e^{-R_0/g(\cdot)} = 1 - c_0 - h_0$ is a transcendental function that cannot be solved algebraically. In the event, we do not need explicit functional forms to sign the derivatives of the choice functions with respect to their ultimate determinants. The choice functions ultimately depend on workers' initial endowment with health, H_0, the constant elasticity of marginal utility, γ, assumed to be greater than one, the consumption elasticity of utility, η_c, and the elasticity of health with respect to health spending, η_h, as well as with respect to retirement age, η_R, as follows:

$R^* = R(H_0, \eta_c, \eta_h, \eta_R)$ with $\partial R^*/\partial H_0 > 0$, $\partial R^*/\partial \eta_c < 0$, $\partial R^*/\partial \eta_h > 0$, and $\partial R^*/\partial \eta_R < 0$;

$h^* = h(H_0, \eta_c, \eta_h, \eta_R)$ with $\partial h^*/\partial H_0 < 0$, $\partial h^*/\partial \eta_c < 0$, $\partial h^*/\partial \eta_h > 0$, and $\partial h^*/\partial \eta_R > 0$;

$c^* = c(H_0, \eta_c, \eta_h, \eta_R)$ with $\partial c^*/\partial H_0 > 0$, $\partial c^*/\partial \eta_c > 0$, $\partial c^*/\partial \eta_h < 0$, and $\partial c^*/\partial \eta_R < 0$.

A higher initial endowment with health, H_0, has a positive effect on the choice of retirement age *R* and on per-period consumption *c*, but a negative effect on health spending *h*. A higher η_c leads to a higher *c* at the expense of *h* and – due to the complementarity between *R* and *h* – also induces earlier retirement, a lower *R*. By contrast, a higher η_h leads to lower *c*, higher *h* and higher *R*. Finally, a higher absolute value of the negative elasticity of health with respect to retirement age, η_R, leads to a lower *R* and a lower *c*, while *h* is increased.

Short term: the indirect utility function for a fixed retirement age

In our model, the value of the optimal adjustment of health spending in the presence of a fixed retirement age R_0 can be inferred from the indirect utility function $\overline{U} = \max H\, u(c)$ whose derivative with respect to η_h is equal to the derivative of the Lagrangian evaluated at the optimal values of the original solution. The Lagrangian for the maximization of $F(x,\alpha) = f(x,\alpha) - \phi(\alpha)$ over the variables in *x* and α, treated as independent variables, is $L^* = f(x,\alpha) -^2\phi(\alpha) + \lambda_1 (g(R_0, h, H_0) - H) + \lambda_2 (1 - e^{-R/g(\cdot)} - c - h) + \lambda_3 (R - R_0)$ with first-order conditions $L^*_x = F_x + \lambda_i\, \varphi_{ix} = 0$, $L^*_\alpha = F_\alpha + \lambda_i\, \varphi_{i\alpha} -^2\phi_\alpha = 0$, and $L^*_{\lambda_i} = \varphi_i = 0$. The first-order conditions from the differentiation with respect to α represent the envelope theorem.

We use the envelope theorem to understand the impact of positive exogenous shocks in healthcare productivity, such as new medical technology. At the optimal choices $x = x^*(\alpha)$, the rate of change of the indirect utility function with respect to η_h, denoted ${}^2\phi_{\eta_h}$, in which only *c* and *h* are allowed to adjust to the change in η_h, is equal to the rate of change of the original Lagrangian with respect to η_h holding R_0 constant. We therefore take the derivative of the Lagrangian with respect to η_h:

$$L^*_{\eta_h} = F_{\eta_h} + \lambda_{i\eta_h}\, \varphi_{i\eta_h} -^2\phi_{\eta_h} = \left[u(\partial g/\partial h) - e^{-R_0/g} (R_0/g) u_c (\partial g/\partial h) - g\, u_c \right] (\partial h/\partial \eta_h).$$

The higher *h* that is induced by the exogenous increase in η_h affects overall utility *positively* through the gain in average per-period utility with better health and *negatively* through the marginal utility impact of the lower per-period income that a fixed *R* implies amid rising life expectancy as well as through the expected loss of marginal utility from consumption amid rising *h*. As in previous sections, the higher average utility of additional lifetime compared with additional consumption is crucial in generating a positive effect overall.

Long term: the indirect utility function with unconstrained retirement timing

Using the indirect utility function ${}^{1}\phi(\alpha)$ without fixed retirement age yields the Lagrangian derivative

$$L_{\eta_h} = F_{\eta_h} + \lambda_{i\eta_h}{}^{1}\varphi_{i\eta_h} - {}^{1}\phi_{\eta_h} =$$

$$\left[u(\partial g/\partial h) - e^{-R_0/g}(R/g)u_c(\partial g/\partial h) - g\,u_c\right](\partial h/\partial\eta_h) + \left[u(\partial g/\partial R) + e^{-R/g}\,u_c\right](\partial R/\partial\eta_h).$$

The first part of this expression is identical to the effect under fixed *R*, the second part adds the impact of increasing *R*. Like before, the consequences of a larger *h* include the utility gain from living longer, the marginal utility loss from lower consumption, due to the lower per-period income when lower mortality is combined with a fixed *R*, and the lifetime marginal utility loss from lower *c*. The consequences of a higher *R* include the loss in average utility from the greater mortality that later retirement induces and the marginal utility from additional consumption that higher lifetime income makes feasible.

Comparison

At the initial values of the exogenous parameters, the fixed-retirement-age constraint is just binding at R_0 and the Lagrange multiplier $\lambda_{3\eta_h}$ is zero, indicating zero marginal cost of the restriction in terms of attainable values of the indirect utility function. The two indirect utility functions have the same value at that point. However, as η_h increases, R_0 becomes strictly binding and reduces the maximum value of the indirect utility function since *R* is no longer available as a choice variable. The constrained short-term indirect utility function is therefore more concave, or less convex, than the indirect utility function without the retirement age constraint. Looking at second-order changes, Taylor series expansions can be used to show that the Hessian matrix of second partials of ${}^{2}\phi(\alpha)$ – ${}^{1}\phi(\alpha)$ is negative semidefinite, which implies that ${}^{2}\phi(\alpha)$ – ${}^{1}\phi(\alpha)$, too, is a concave function of the exogenous parameters.

Illustration

We now illustrate the complementarity of *h* and *R* in the production of health by means of a simple graph of intertemporal health transitions. Stable fixed points in these transitions correspond to equilibria identified in our formal model. The underlying idea is that we interpret the indirect utility function as the value function of a dynamic programming problem, which can then be expressed in terms of the familiar Bellman equation. To this end, we divide the value function

$$V(g) = \max_{c,h,R} g(R,h)\, u(c) \quad \text{s.t.} \quad e^{-R/g} = 1 - c - h$$

by $g(R,h)$ and obtain the Bellman equation with discounting only for mortality, not for time preference,

$$V(g) = \max_{c,h,R} u(c) + [1 - 1/g(R,h)]V(g) \quad \text{s.t.} \quad e^{-R/g} = 1 - c - h\,.$$

In the budget constraint, $e^{-R/g(\cdot)} = 1 - c - h$, the left-hand side will be smaller the larger *R*; either *c* or *h* – or both to a smaller degree – can hence be increased as *R* increases. However, a higher *R* also lowers $g(R,h)$ and thus raises the marginal product of *h* so that *h* must be increased as part of an optimal policy.

Figure 3 illustrates the dynamic programming approach in terms of a health transition function that would be plausible under conditions of work-related health deterioration. Consider first the lower of the two curves, starting in the origin and passing through points *B* and *A*. Because the marginal efficiency of health spending is assumed to decline, maintaining health at its maximum value of one would not be optimal. Instead, we assume a fixed point at *A* that is locally stable because any health shocks that do not move the worker's health below *B* can still be partially repaired from one period to the next. However, when a worker's health falls below *B*, recovery of health is no longer feasible while the worker continues to work and continued work will lead to death, as indicated by the stable fixed point at zero.

With the upper curve that passes through points *D* and *E*, Figure 3 illustrates a new health transition function for a *retiring* worker who retains access to medical care. Because some of the health shocks previously suffered from work may be irreparable while working, but have no lasting effect in retirement, the upper stable fixed point moves from *A* to *E*. Moreover, the minimum level of health from which recovery is still feasible moves from *B* to *D*. Finally, we assume there is a second stable fixed point at *C* that avoids death. This accounts for the observation that there are many diseases that have a permanent adverse effect on health but need not be fatal if the person receives proper cure and care in full retirement from work.

How would an exogenous increase in η_h, such as a boost in healthcare productivity from new medical technology, change the picture? In terms of Figure 3, a higher η_h would move point *A* to the right and point *B* to the left, narrowing or eliminating the intervals on the H_t-axis from which a sick worker must resort to retiring early in order to reach the stable fixed point with good health, point *E*.

V. DISCUSSION AND CONCLUSION

The management of healthcare amid population aging is a focal challenge for the world economy in the 21st century. Two issues that are interrelated require particular attention: financing the rising demand for healthcare and investing in the appropriate medical technologies. Solving the financial side requires anticipating the opportunities that new technologies create. Setting the right priorities for technology investments requires anticipating the financial constraints and rewards. Hall and Jones (2007, p. 30) predict that "maximizing social welfare in the United States will require the development of institutions that are consistent with spending 30 percent or more of GDP on health by the middle of the century." We have argued that these institutions should include transparent mechanisms to coordinate increases in health spending and retirement age.

A key question in this context – how much can the health of the elderly be expected to improve – is an empirical one. It is often framed in terms of two competing hypotheses, known as morbidity expansion and morbidity compression, and a synthesis in terms of a dynamic equilibrium. Under morbidity *expansion*, people will spend a rising share of their total lifetime in a state of chronic morbidity, often suffering from multiple degenerative diseases that defy a comprehensive diagnosis and cure, treated with drugs that merely keep patients alive at a reduced quality of life. Morbidity *compression* is more optimistic and predicts that future longevity gains will be accompanied by a compression of sick time in the vicinity of death. Population aging would hence not force lifecycle per-capita health spending to rise very much. The *dynamic equilibrium* hypothesis synthesizes these extreme predictions: Rising longevity will allow the average lifetime *share* that people spend in morbidity to fall, even if total sick time before death does not shrink in absolute terms. We think that both the morbidity compression and the dynamic equilibrium hypotheses are consistent with large welfare gains from the coordinated postponement of retirement and a rising health share in income, provided the growth in spending helps to further reduce morbidity and mortality among the elderly. The type of new medical technology and the direction of biomedical research may hence also become legitimate targets for government-sponsored coordination.

How can such coordination be successful and why has it not happened before? In some sense, the government's coordination problem is analytically similar to the case of policy complementarities in fighting European unemployment in Coe and Snower (1997). They define a group of policies as complementary when the effect of each policy on the target is greater through joint implementation with the other policies than in isolation. Governments may fail to realize and exploit the relevant complementarities for a multitude of reasons. One *proximate* reason is that health and retirement policies are often designed in different departments and decided upon at different points in times. The responsible policy makers simply may not see the complementarities and opportunities for cooperation. Moreover, those with the necessary expertise and influence in one system may suffer from regulatory capture or may face a prisoner's dilemma: Postponing the mandatory retirement age unilaterally without substantial investments in workers' health is neither likely to be efficient, nor acceptable on distributional grounds. On the other hand, much higher health spending without extending people's working lives is unlikely to be fiscally sustainable. This dilemma is compounded by the difficulty of any democratic government to make credible commitments over the decades-long span of time between new medical technology investments and their full social payoff.

Even if these problems can be overcome, governments may still be constrained by a lack of good scientific evidence about empirical linkages between health expenditures, health status and welfare, as pointed out by Dormont et al. (2007), by capital market imperfections of the kind emphasized by Kalemli-Ozcan and Weil (2002), Philipson and Becker (1998), Cremer et al. (2004) and Dow et al. (1999), and by a critical lack of international coordination. Many countries have aging populations and the markets for medical technology have long become global. Efficient innovation incentives thus require a global strategy of aggregating consumers' willingness-to-pay in all countries likely to benefit, notwithstanding differences in preferences they may have. We do not yet have a solution for this.

REFERENCES

Becker, G.S., T.S. Philipson and R.S. Soares (2005). The Quantity and Quality of Life and the Evolution of World Inequality. *American Economic Review* 95(1): 277–291.

Coe, D.T. and D.J. Snower (1997). Policy Complementarities: The Case for Fundamental Labor Market Reform. *IMF Staff Papers* 44(1): 1–35.

Cremer, H., J.-M. Lozachmeur and P. Pestieau (2004). Retirement Age and Health Expenditures. CEPR Discussion Paper 4774.

Cutler, D., A. Deaton and A. Lleras-Muney (2006a). The Determinants of Mortality. *Journal of Economic Perspectives* 20(3): 97–120.

Diamond, P.A. (1965). National Debt in a Neoclassical Growth Model. *American Economic Review* 55(5): 1126–1150.

Dormont, B., J. Oliveira Martins, F. Pelgrin and M. Surcke (2007). Health Expenditure, Longevity and Growth. Paper presented at IX European Conference on the Fondazione Rodolfo de Benedetti "Health, Ageing and Productivity", Limone sul Garda.

Dow, W.H., T.J. Philipson, and X. Sala-i-Martin (1999). Longevity Complementarities under Competing Risks. *American Economic Review* 89(5): 1358–1371.

Dwyer, D.S. and O.S. Mitchell (1999). Health Problems as Determinants of Retirement: Are Self-rated measures Endogenous? *Journal of Health Economics* 18: 173–193.

European Commission (2006). The Impact of Aging on Public Expenditure: Projections for the EU-25 Member States on Pensions, Healthcare, Long-term Care, Education and Unemployment Transfers (2004-50). European Economy Special Report 1/2006.

Gruber, J. and D. Wise (1998). Social Security and Retirement: An International Comparison. *American Economic Review* 88(2): 158–163.

Hall, R.E. and C.I. Jones (2007). The Value of Life and the Rise in Health Spending. *Quarterly Journal of Economics* 122(1): 39–71.

Hostenkamp, G. and M. Stolpe (2008). The Social Costs of Health-related Early Retirement in Germany: Evidence from the German Socio-economic Panel. Kiel Working Paper 1415.

Kalemli-Ozcan, S. and D.N. Weil (2002). Mortality Change, the Uncertainty Effect, and Retirement. NBER Working Paper 8742.

Nordhaus, W.D. (2003). The Health of Nations: The Contribution of Improved Health to Living Standards. In: K.M. Murphy and R. Topel (eds.), *Measuring the Gains from Medical Research: An Economic Approach*. Chicago and London.

Oeppen, J. and J.W. Vaupel (2002). Broken Limits to Life Expectancy. *Science* 296(5570): 1029–1031.

Philipson, T.J. and G.S. Becker (1998). Old-Age Longevity and Mortality-Contingent Claims. *Journal of Political Economy* 106(3): 551–573.

Queisser, M. and E. Whitehouse (2006). Comparing the Pension Promises of 30 OECD Countries. *International Social Security Review* 59(3): 49–77.

Rosen, S. (1988). The Value of Changes in Life Expectancy. *Journal of Risk and Uncertainty* 1: 285–304.

Sala-i-Martin, X. (1996). A Positive Theory of Social Security. *Journal of Economic Growth* 1(2): 277–304.

Samuelson P.A. (1958). An Exact Consumption-loan Model of Interest With or Without the Social Contrivance of Money. *Journal of Political Economy* 66: 467–482.

Silberberg, E. (1990). The Structure of Economics: A Mathematical Analysis. Second Edition. New York.

Westerhout E. and F. Pellikaan (2005). Can We Afford to Live Longer in Better Health? CPB Document 85.

Weil, D.N. (2007). Accounting for the Effect of Health on Economic Growth. *Quarterly Journal of Economics* 122(3): 1265–1306.

FIGURES

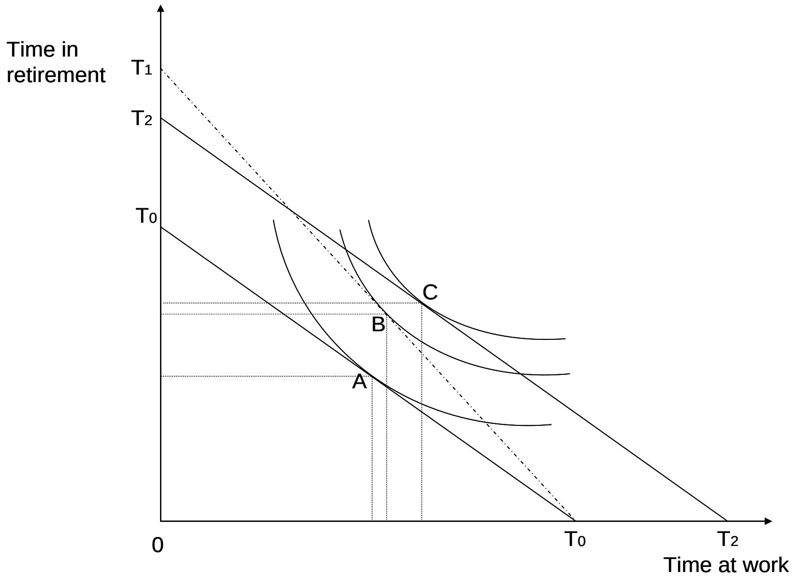

Figure 1:
Price and Income Effects in Retirement Timing

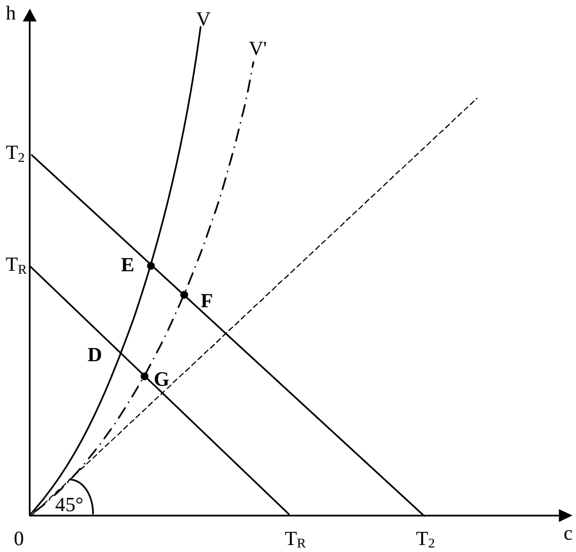

Figure 2:
The Static Model with Endogenous Health Spending (V) and Retirement Timing (V')

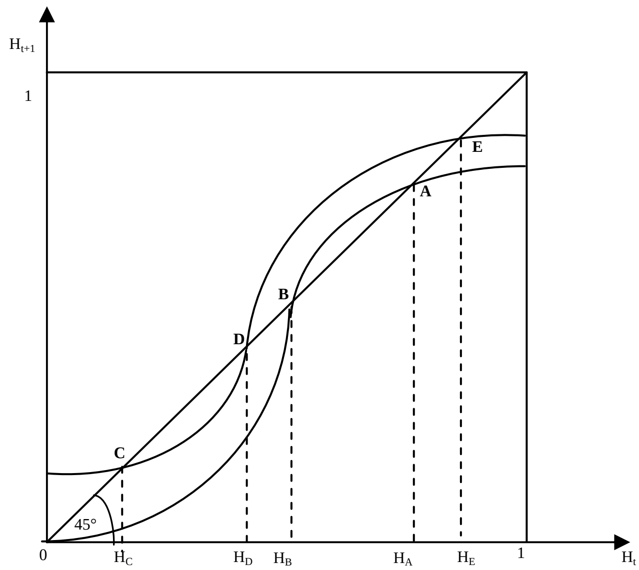

Figure 3:
Health Transition Functions with Work-related Health Deterioration (lower curve) and the Early Retirement Option (upper curve)