DIW BERLIN

1226

Discussion Papers

Health Care Expenditures and Longevity: Is there a Eubie Blake Effect?

Friedrich Breyer, Nomann Lorenz and Thomas Niebel

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

Health Care Expenditures and Longevity:
Is there a Eubie Blake Effect?

Friedrich Breyer[*], Normann Lorenz[†] and Thomas Niebel[‡]

July 12, 2012

Abstract

It is still an open question whether increasing life expectancy as such is causing higher health care expenditures (HCE) in a population. According to the "red herring" hypothesis, the positive correlation between age and HCE is exclusively due to the fact that mortality rises with age and a large share of HCE is caused by proximity to death. As a consequence, rising longevity – through falling mortality rates – may even reduce HCE. However, a weakness of previous empirical studies is that they use cross-sectional evidence to make inferences on a development over time. In this paper we analyse the impact of rising longevity on the trend of HCE over time by using data for a pseudo-panel of German sickness fund members over the period 1997-2009. Using (dynamic) panel data models, we find that age, mortality and five-year survival rates have a positive impact on per-capita HCE. Our explanation for the last finding is that physicians treat patients more aggressively if they think the result will pay off for a longer time span, which we call "Eubie Blake effect". A simulation on the basis of an official population forecast for Germany is used to isolate the effect of demographic ageing on real per-capita HCE over the next decades.

JEL-classification: H51, J11, I19.

Keywords: health care expenditures, ageing, longevity, 5-year survival rate.

If I'd known I was going to live this long,
I would have taken better care of myself.
(Eubie Blake on his alleged 100th birthday)

1   Introduction

The ageing of population in most OECD countries will place an enormous burden on tax payers over the next decades. Given this demographic change, previous fiscal policies in several of these countries were unsustainable, and major reforms of social insurance systems have been enacted, in particular with respect to public pension and long-term care financing systems. However, what remains unclear is whether population ageing also jeopardizes the sustainability of social health insurance (see, e.g. Hagist and Kotlikoff (2005) and Hagist et al. (2005)). While there is no doubt that the revenue side of these systems will suffer from the shrinking size of future taxpayer generations, it is not so clear if rising longevity will place an extra burden on the expenditure side. If so, additional reforms of these systems would be necessary to guarantee the sustainability of these systems, such as introducing more funding or limiting the generosity of benefits.

The impact of population ageing on health care expenditures (henceforth: HCE) has been heavily debated over the last decade.[1] That a positive association of age and HCE is primarily due to the high cost of dying and rising mortality rates with age, was first observed by Fuchs (1984). Subsequently, Zweifel et al. (1999) have coined the term "red herring" to character-ize the erroneous conclusion from this cross-section correlation that population ageing due to increasing longevity implies rising country level HCE over time. As counter-evidence they showed that in individual data – when controlling for proximity to death – calendar age is not even a significant predictor of health care costs.

Although this early study suffered from its focus on patients in their last year of life, sub-sequent studies by several authors such as Stearns and Norton (2004), Seshamani and Gray (2004), Zweifel at al. (2004) and Werblow et al. (2007) confirmed the red herring hypothe-sis by demonstrating that even for persons who survived for at least four more years, there is hardly any age gradient in HCE, whereas the costs of the last year of life tend to decrease with the age at death (Lubitz et al. (1995)). The latter finding is explained by the tendency of physicians to treat patients who have lived beyond a "normal life-span" less aggressively than younger patients with the same diagnosis and the same survival chances. In this vein, Miller (2001) shows by simulation that, based on a negative relationship between age at death and death-related costs, an increase in longevity will dampen the growth of HCE.

However, an important weakness of almost all studies in the related literature is their reliance on cross-section expenditure data. Therefore, in drawing inferences from these studies for the development of HCE over time, proponents of the red herring hypothesis commit the same error of which they accuse their opponents (i.e. those who believe that population ageing in-creases health spending because per-capita expenditures increase with age). In particular, they overlook the fact that increasing longevity not only means that 30 years from now average age at death will be higher, but also that people at a certain age (say, 80) will on average have more years to live than present 80-year olds.

It is reasonable to assume that physicians, who have to allocate scarce resources among their patients, will make a conjecture how long a patient will benefit from a treatment (such as implanting an artificial hip), and this depends upon the patient's expected longevity. In that

---

[1] A recent survey can be found in Karlsson and Klohn (2011).

respect, the physician will behave in a way described in the famous quotation by Eubie Blake, i.e. he will spend more on patients who will profit from his treatment for a longer time period.[2]

This effect will lead to a similar physician behaviour as "age-based rationing" of health care services when the notion of a "normal life span" (Callahan (1987), Daniels (1985)) shifts over time with rising longevity. Indeed, the empirical literature shows that some physicians use age as a prioritization criterion in allocating scarce health care resources (for an overview see Strech et al. (2008)).

This reasoning suggests that the relationship between "life expectancy" or "time to death" and HCE is non-monotonous: in the very last years of life, a *lower* value of these variables indicates worse health and therefore higher HCE, e.g. for emergency treatment and heroic efforts to avoid the unavoidable. In individual data, this effect can be captured by a dummy for "last year of life" and in group data by the share of persons who died in the particular year, i.e. the mortality rate. In contrast, when time to death is longer (say, between 5 and 10 years), a *higher* value indicates a better chance to benefit from elective surgery and other potentially risky procedures for a longer time and thus leads to higher HCE, as argued above. This "Eubie Blake effect" can be captured in group data by including a measure for longevity, holding the mortality rate constant.

To test whether there is a "Eubie Blake effect", it is desirable to study how rising life expectancy in a population has affected health care expenditures *over time*. This requires a data set that comprises this variable, or an indicator of it, and covers several years.

To our knowledge, there have been only three previous studies which have used life expectancy as an explanatory variable in a regression equation for HCE, viz. Shang and Goldman (2008), Zweifel et al. (2005) and Bech et al. (2011), of which the first one used individual-level data, the other two population-level data.

Shang and Goldman (2008) used a rotating panel of more than 80,000 Medicare beneficiaries and predicted the life expectancy for each individual, based on age, sex, race, education and health status and then performed a nonlinear-least-squares estimation of individual HCE. In this equation, predicted life expectancy turned out to be highly significant and negative, whereas age became insignificant when this variable was included. The interpretation of this result is, however, very similar to other studies in the red herring literature because predicted life expectancy, if the value is low (say, a few years), is a proxy for time to death.

Zweifel et al. (2005), in contrast, used a panel of 17 OECD countries over a period of 30 years (1970-2000) and tried to jointly explain HCE and life expectancy. As one of the determinants of HCE, they constructed an artificial variable "SISYPH" (for Sisyphus effect) by multiplying "life expectancy at 60" (averaged over both sexes) with the share of persons over 65 in the total population. The predicted value of this variable turned out to be a significantly positive predictor of HCE. A problem with this result is that it does not allow disentangling the effects of the old age dependency ratio and life expectancy itself.

Bech et al. (2011) considered per-capita HCE for a panel of 15 EU member states over the period 1980 to 2003 and found that both mortality and remaining life expectancy at age 65 have a significant positive effect on HCE in the following year. They then calculated long-run elasticities of HCE with respect to these variables and found a positive value only for life expectancy, so that a linear increase in life expectancy at 65 is associated with an exponential growth in per-capita HCE.

---

[2]Fang et al. (2008) attribute the same quotation to the baseball star Mickey Mantle and speak of a "Mickey Mantle effect". However, it is quite clear that Mantle did not invent the sentence, but quoted the football player Bobby Lane, who died in late 1986 and may well have known the statement by Blake, which was made already in February 1983.

In this paper, we aim at disentangling the two effects of rising longevity, i.e. the "direct" effect of decreasing HCE (at a certain age) due to a falling mortality rate (at that age) and the "indirect" effect of increasing HCE due to an increase in the remaining life expectancy at that age (conditional on surviving until the end of the year). To do so we employ a measure for remaining life expectancy, which is especially common among physicians: (expected) 5-year survival rates. In medical studies, in particular those concerned with specific diseases, this measure is used instead of life expectancy as such.

The data set we employ is a pseudo panel of sickness fund members in Germany, which was originally collected for calculating age and sex specific (average) HCE for purposes of risk adjustment. This data set, which covers the years 1997 to 2009, is merged with data on mortality rates published annually by the Human Mortality Database. We are convinced that population-level data are suitable for estimating such an effect because the controversy on the red herring effect – even though many authors use individual data to make their point in this debate – is focused exactly on the question whether population ageing will lead to increasing HCE *in a country*, thus jeopardizing the fiscal soundness of publicly financed health systems.

To determine the impact of longevity we estimate (dynamic) panel data models; to disentangle age, period and cohort effects, we apply the Intrinsic Estimator (Yang et al. (2008)), which is a special case of a partial least squares regression (Tu et al. (2012)). We then use the estimated relationship to show the effect of an increase in survival rates according to official statistics on average HCE. We find that while falling mortality rates as such lower HCE, this effect is more than compensated by an increase in remaining life expectancy so that the net effect of ageing on HCE over time is clearly positive.

The remainder of this paper is organized as follows. In Section 2 we describe the data, in Section 3 we state the theoretical hypotheses to be tested, in Section 4 we explain the methodology of estimating the determinants of HCE, in Section 5 we present the regression results, in Section 6 we perform a simulation of the future development of HCE, and Section 7 concludes.

2 Data

The data used in this study come from three different sources. Data on HCE were provided by the German Federal (Social) Insurance Office ("Bundesversicherungsamt", BVA). To determine the risk adjustment payments for the statutory sickness funds, each year the BVA collects data on all expenditures covered by the sickness funds for all individuals insured in the social health insurance system. These data comprise eight major expenditure categories including inpatient care, ambulatory care, dental care and prescription drugs. Based on this census, the BVA calculates and publishes the average HCE for all sickness fund members, separately for each age-sex group.[3] It also publishes the number of individuals in each group.[4]

Data on age and sex specific mortality rates are taken from the Human Mortality Database (2011), and were used to calculate 5-year survival rates. These data apply to the German population as a whole and not only to sickness fund members. Since the omitted group, the privately insured, have on average higher incomes, and life expectancy is positively associated with income in Germany (von Gaudecker and Scholz (2007), Breyer and Hupfeld (2009)), the

---

[3]The official risk adjustment data, which the BVA publishes on its website, are smoothed. We use the unsmoothed data and thank Dirk Göpffarth, the Head of the Risk Adjustment Unit at BVA for making this data set available to us.

[4]To be more precise, the variables are *average HCE per day* and *number of person-days*, i.e., the number of insured times the average number of days per year an individual of this age-sex group is insured. In addition, the data set contains these two variables also separately for the two regions East and West Germany, however only until 2007. Since 2008 there is no distinction according to region in the risk adjustment scheme any more.

population-based survival rates are somewhat higher than the true survival rates of sickness fund members, but this error should be rather small given that sickness fund members account for about 90 per cent of the German population.

In the HCE data set provided by the BVA, the highest age group contains the average HCE of all individuals of age 90 and above. Since we have no information about the age distribution within this group, we could not compute their average mortality and survival rate. We therefore drop this group, which amounts to a loss of 0.71% of all person-days.

Our data set comprises the period 1997 to 2009. As there are 90 age groups (0 to 89) for men and women separately, the total number of observations is 2340. Table 1 contains descriptive statistics on the data set. Since we perform the estimations separately for men and women, we present these statistics separately, too. For men, average HCE per day range from € 1.78 (at age 3 in 1997) to € 17.60 (at age 89 in 2009).

Table 1: Descriptive Statistics of the Data Set

	Men				**Women**			
	mean	std.dev.	min	max	mean	std.dev.	min	max
Age	44.5		0	89	44.5		0	89
Cohort	1958.5		1908	2009	1958.5		1908	2009
HCE	6.2437	4.7329	1.7812	17.6005	6.1312	3.8728	1.5020	15.7070
$MORT$	.0233	.0437	.00007	.2275	.0153	.0321	.00005	.1711
$SR5$	.8785	.2021	.1687	.9996	.9117	.1685	.2603	.9997

Table 2 presents 5-year survival rates for selected age groups in the base year 1997 and their increase over time until 2009. The table shows that there is a great amount of variation in these rates both over time and across age groups. In particular, between 1997 and 2009, 5-year survival rates have been increasing by up to 9 percentage points for men; for women the increase is smaller but still up to 5.6 percentage points.

Table 2: 5-year survival rates: Level in 1997 (per cent) and increase Δ from 1997 to 2009 (percentage points)

	Men		**Women**	
	$SR5$		$SR5$	
Age	1997	Δ	1997	Δ
60	91.1	2.4	95.9	0.8
65	86.1	4.3	93.2	1.9
70	79.1	5.9	88.3	3.4
75	67.9	6.9	79.5	4.6
80	51.2	9.0	64.6	5.6
85	31.6	8.6	43.6	4.7
90	14.0	4.0	22.1	1.1

The third source of data for our study is the German Statistical Office. Every three years, it publishes forecasts on the size and composition of the population in Germany over the following decades. The most recent one is the "12th coordinated population projection" (Statistisches Bundesamt 2009). In addition, the Office provided estimates of the development of age-specific mortality rates over the period until 2060. From these data, we calculated the time paths of age-specific survival rates. Of the two published forecasts, the one denoted the "most likely one" by the Office and the one with an even stronger increase in longevity, we use the "most likely one". These data were used for simulating the demographic effect on HCE in Section 6.

3   Testable Hypotheses

The main focus of the paper will be the effect of "population ageing", expressed by falling mortality rates and increasing life expectancy, on average HCE of a population group. However, age and time will be used as explanatory variables in the regression as well. The following theoretical predictions are derived from the literature and will be tested in the empirical estimation:

Age: According to more "traditional" theory, HCE will be decreasing with age in the age range 0-20, approximately constant between 20 and 60 and increasing with age for age above 60. In contrast, the alternative hypothesis on which the red herring claim is based states that HCE will be independent of age for age above 20.

Time: HCE will be increasing over time due to medical progress.

Mortality: As for individuals expenditures are especially high in the last year of life, average HCE of a population group will be increasing in the mortality rate ($MORT$) of the group.

Life expectancy: Holding the mortality rate of an age group constant, HCE of this group will be increasing in the remaining life expectancy within the group as physicians will spend more resources on patients who have "more to gain" from an intervention. This "Eubie Blake effect" is especially important for older patients.

Throughout this paper, "life expectancy" will be measured by the 5-year survival rate $SR5$ (of each age group), which is a familiar concept for physicians. Using the 5-year survival rate has an important advantage over the variable "life expectancy" itself: In younger age groups life expectancy falls almost linearly with age, whereas 5-year survival rates vary very little with age and start falling only later. Here the variance with age and over time occurs almost exclusively in older age groups, and the effect of this on HCE is exactly what we want to test. We note that technically the 5-year survival rate at age a in year t conditional on surviving at least until the end of year t is calculated by multiplying the one-year survival rates (i.e. one minus the mortality rate) of age groups $a+1$, $a+2$, ..., $a+5$ in year t. This corresponds to the usual way remaining life expectancy for an age group is calculated.

For the following reason we do not use the 5-year survival rate $SR5$ as such but a predicted value of it. We argued that a physician will take the 5-year survival rate into account when deciding whether to perform an expensive or risky procedure or on which patients to ration (most). However, during the year t, the physician does not know the 5-year survival rate $SR5_{c,a,t}$, as this is a measure derived from the mortality rates in the same year, which are not known until the end of the year. It is therefore an informed guess of the survival rate the physician will have in mind. One possible proxy for this guess would be the value of this

variable in the previous year (for the same age), $SR5_{c-1,a,t-1}$, but this is certainly not the best choice: First, survival rates are increasing over time, so there would be a systematic downward bias in this proxy. Secondly, as the survival rate in a particular year $t-1$ is derived from the mortality rates of 5 different age groups in $t-1$, they may depend heavily upon singular events such as a flu epidemic or a heat wave. These singular events will however have no (or only a minor) effect on the informed guess of the physician. Rather, it will depend on his or her experience over a longer time period, so we use a linear projection of the survival rate (for the same age) of the previous five years.[5] In the following, whenever we use the symbol $SR5$, we refer to this prediction of the 5-year survival rate.[6]

4 Estimation Strategy

Before we present the estimation strategy, we give an overview of how we denote the different variables used in the regressions:

- $HCE_{c,a,t}$ (dependent variable), the average value of daily health care expenditures of all insured persons in cohort c of age a in year t, converted to Euros of 2009 by using the consumer price index;

- a set of dummy variables Age_a for each age a with $a = 0, \ldots, 89$;

- a set of dummy variables $Cohort_c$ for each cohort c with $c = 1908, \ldots, 2009$, (the year in which the person was born);

- a set of dummy variables $Year_t$ for each year t with $t = 1997, \ldots, 2009$;

- $MORT_{c,a,t}$, the mortality rate, i.e. the share of persons in cohort c of age a in year t who die within that year;

- $SR5_{c,a,t}$, the predicted 5-year survival rate of all persons in cohort c of age a in year t.

As each entry in HCE, $MORT$ and $SR5$ contains the average values of a particular age-sex-group, our data set is a "pseudo panel" in the sense of Deaton (1985). For these pseudo panels, Verbeek and Nijman (1992) have shown that for a sufficiently large number of individuals in each group, the group averages are unbiased estimators of the "true" value in the population.

To describe the estimation strategy, we begin with the general specification

$$HCE_{c,a,t} = g(c, a, t) + \beta_1 MORT_{c,a,t} + \beta_2 SR5_{c,a,t} + u_{c,a,t}, \tag{1}$$

where g captures the effects of cohort, age and time, and $u_{c,a,t}$ denotes the error term. There is no dummy variable included for sex, because we perform all estimations separately for men

---

[5]Technically, we run a regression of $(SR5_{c-5,a,t-5}, \ldots, SR5_{c-1,a,t-1})'$ on a constant and a linear time trend, i.e.

$$SR5_{c-\tau,a,t-\tau} = \mu_0 - \mu_1\tau \quad \text{for} \quad \tau = 1, \ldots, 5,$$

and determine the prediction as $\widehat{SR5}_{c,a,t} = \hat{\mu}_0 - \hat{\mu}_1 \cdot 0 = \hat{\mu}_0$. For each prediction a separate regression is performed. As these equations are estimated for every age, there is implicitly an interaction between age and year in this estimation. In contrast, there is no age-year interaction term in the equation for HCE, which solves the identification problem.

[6]One might be concerned that because $MORT$ and $SR5$ are highly correlated the effect of the two variables cannot be disentangled empirically. The regression results reported in Section 5, however, show that this concern is not justified: standard errors are quite small and the coefficients of both variables for men and at least of $SR5$ for women are (highly) significant in all models.

and women, since – as it is well known – the age profiles of HCE have rather different shapes for men and women.

The specification in (1) suffers from the familiar problem of perfect multicollinearity since age equals year minus cohort:

$$a = t - c. \tag{2}$$

Because we want to estimate the effects of cohort, age and time in a flexible manner, we follow the dummy-variables approach and set

$$g(c, a, t) = \beta_0 + \sum_c \gamma_c Cohort_c + \sum_a \alpha_a Age_a + \sum_t \delta_t Year_t, \tag{3}$$

where in each set of the dummy variables one variable is skipped because of the constant term.

Of course, the problem of perfect multicollinearity applies to the dummy variables specification as well. There are in principle two strategies to deal with this problem: The first one is to drop one of the variables (or set of dummies for) age, cohort or time and for example estimate a model with only age- and year-dummies. Because our data set is a pseudo panel where the "individuals" are cohorts, this variable cannot be dropped in our analysis. Obviously, neither the age effect nor the year effect (medical progress) can be dropped, either.

The second strategy then is to impose a restriction on the coefficients γ, α and δ.[7] One can distinguish two ways to do so: In most cases, one of the coefficients is set to zero, or two – usually but not necessarily adjacent – coefficients are set equal. E.g. with $\delta_{2000} = \delta_{2001}$, it is assumed that there is no time effect going from the year 2000 to 2001; with $\alpha_{20} = \alpha_{21}$, it is assumed that 20 and 21-year-olds have equal health care expenditures. If one can be confident that the assumption is valid, this will correctly disentangle the age, period and cohort effects.

However, as shown by Yang et al. (2008), the resulting estimates can be seriously misleading, if the assumption is not warranted. In fact, in our data the estimates are very sensitive to which two coefficients are set equal: If, for example, we assume $\alpha_{23} = \alpha_{24}$, the year dummies indicate a positive time trend; this reverses for $\alpha_{24} = \alpha_{25}$, so that HCE are estimated to decrease over time. For $\alpha_{25} = \alpha_{26}$, the time trend is again positive.[8] This lack of robustness is a strong reason for discarding this solution to the multicollinearity problem.

The second way to impose a restriction on γ, α and δ is the following: The problem in estimating (1) with $g(a, c, t)$ replaced by the set of dummies as shown in (3) is that the well known least squares formula $(X'X)^{-1}X'HCE$, (where X is the matrix containing all the explanatory variables) cannot be applied because $X'X$ is a singular matrix that cannot be inverted. However, an infinite number of generalized inverse matrices exist. A particular one is the Moore-Penrose inverse, which, as Tu et al. (2012) point out, is to be preferred, because it minimizes the variance of the least squares solution. More importantly, they show that the results using the Moore-Penrose inverse correspond to the results of both a Principal Component Regression and a Partial Least Squares Regression if the maximum number of components is used, and also coincides with the Intrinsic Estimator proposed by Yang et al. (2008).[9]

---

[7] Of course, dropping one of the variables means imposing the restriction that all coefficients on this variable are zero. However, since this is usually not made explicit, we mention it as a separate way to deal with the problem of perfect multicollinearity.

[8] The time trend is estimated to be negative if the coefficients are set equal for ages 6/7, 7/8, 13/14, 14/15, 15/16, 16/17 or 24/25 for men, and similarly for women.

[9] However, because of the way the Intrinsic Estimator is implemented in STATA's apc_ie command, in practice the estimates of the Intrinsic Estimator may differ somewhat. In our data, we find the difference between the estimates using the Partial Least Squares Regression procedure of the software package R and STATA's apc_ie command to be

The restriction on the coefficients γ, α and δ that is effectively imposed by these methods is that the effect of age and cohort is set equal to the effect of time, (for the following exposition see Tu et al. (2012)): With the continuous variables a, c and t for age, cohort and time, the restriction inherent in the variables is $c + a = t$; the same restriction is then imposed on the coefficients: $\gamma + \alpha = \delta$.

The equivalent restriction inherent in the dummy variables is

$$\sum_a \left(a - \frac{A+1}{2} \right) Age_a + \sum_c \left(c - \frac{A+T}{2} \right) Cohort_c = \sum_t \left(t - \frac{T+1}{2} \right) Year_t, \quad (4)$$

(where A is the number age dummies and T the number of year dummies), which, when translated to the coefficients yields:

$$\sum_a \left(a - \frac{A+1}{2} \right) \alpha_a + \sum_c \left(c - \frac{A+T}{2} \right) \gamma_c = \sum_t \left(t - \frac{T+1}{2} \right) \delta_t. \quad (5)$$

Instead of imposing the restriction that one of the coefficients is zero, or two coefficients are equal, the constraint is in a sense "spread more evenly over all coefficients." This, as already mentioned, minimizes the variance of the least squares solution.

Before we proceed, it is important to note that the restriction imposed on γ, δ and α has no influence on the coefficients of all the other covariates: Regardless of whether one coefficient is set equal to zero, or two coefficients are set equal, or the Intrinsic Estimator is used, $\hat{\beta}_1$ and $\hat{\beta}_2$ will always be the same. That means that the coefficients we are most interested in are not at all affected by how the collinearity problem is solved. As a consequence, the predicted values $\widehat{HCE}$ also do not depend on which restriction is imposed.[10]

To have a comparison model as it is used by the proponents of the red herring hypothesis we first estimate (using the Intrinsic Estimator)

$$HCE_{c,a,t} = \beta_0 + \sum_c \gamma_c Cohort_c + \sum_a \alpha_a Age_a + \sum_t \delta_t Year_t + \beta_1 MORT_{c,a,t} + u_{c,a,t}. \quad (6)$$

We than add the five-year survival rate to estimate (again using the Intrinsic Estimator)

$$\begin{aligned} HCE_{c,a,t} &= \beta_0 + \sum_c \gamma_c Cohort_c + \sum_a \alpha_a Age_a + \sum_t \delta_t Year_t + \beta_1 MORT_{c,a,t} \\ &\quad + \beta_2 SR5_{c,a,t} + u_{c,a,t}. \end{aligned} \quad (7)$$

There are two possible reasons why (7) may still be misspecified: First, the true relationship may be dynamic so that there is persistence in HCE. To account for this problem, we also estimate the following dynamic panel model:

$$\begin{aligned} HCE_{c,a,t} &= \phi HCE_{c,a-1,t-1} + \gamma_1 MORT_{c,a,t} + \gamma_2 SR5_{c,a,t} \\ &\quad + \sum_c \beta_c Cohort_c + \sum_a \alpha_a Age_a + \sum_t \delta_t Year_t + u_{c,a,t}. \end{aligned} \quad (8)$$

negligible. Since we perform the other regressions in STATA, we used the Intrinsic Estimator.

[10]This is an application of the Frisch-Waugh-Lowell-Theorem, see Davidson and MacKinnon (1993), Chapter 1. No matter which restriction is imposed on γ, δ and α, the subspace spanned by (3) is always the same.

Secondly, the variables may be non-stationary so that there may be the problem of spurious regression. For this reason we tested for unit roots. Since these tests do not reject non-stationarity in the explanatory variables, (although they do so for HCE), we also estimate the models (6) to (8) in first (and second) differences, i.e. we replace HCE, $MORT$ and $SR5$ by ΔHCE, $\Delta MORT$ and $\Delta SR5$, (and $\Delta^2 HCE$, $\Delta^2 MORT$ and $\Delta^2 SR5$).

We estimate the dynamic panel model (8) by GMM, using both the difference-GMM-estimator by Arellano and Bond (1991) and the system-GMM-estimator by Blundell and Bond (1998), and show the results both for $MORT$ to be either predetermined or endogenous. We present this rather large number of regression results for the dynamic panel data models to show that the familiar problem of unstable coefficients – depending heavily (even in their sign) on the particular specification – does not apply in our data.

As the data set is a pseudo panel, and the respective cohort-age cells contain different numbers of observations, the results from the simple fixed-effects panel estimation may not be efficient and have to be weighted by the square root of the cohort size, see Deaton (1985). Because in our panel the cohort size is not constant over time, we could use different weights for each cohort-age cell. However, Inkmann et al. (1998) show that estimation results can be unstable if the cohort size differs considerably and therefore propose to weight by the average weight for each cohort. We therefore use weights that do not differ in the time dimension.

5 Regression Results

5.1 Unit root tests

We first employ the unit root tests by Harris and Tzavalis (1999) and by Im et al. (2003) without and with different numbers of lags. Table 3 shows an overview of the results; the detailed results can be found in Tables 8 to 10 in the Appendix. For the dependent variable HCE, non-stationarity is clearly rejected. For $MORT$ and $SR5$, non-stationarity in levels is never rejected, as all p-values are very close to 1. For first differences, the results are ambiguous as the null hypothesis is only rejected for some of the tests. For second differences, the null is always rejected. Therefore, we not only present the results for the estimation in levels, but as a robustness check in first and second differences as well.

Table 3: Unit root tests: Rejection of H_0: non-stationarity

	Men			Women		
	level	Δ	Δ^2	level	Δ	Δ^2
HCE	yes	yes	yes	yes	yes	yes
$MORT$	no	yes/no	yes	no	yes/no	yes
$SR5$	no	yes/no	yes	no	yes/no	yes

5.2 Estimation results

In Tables 4 and 5 we present the regression results, separately for men and women. In column (1), results from the Intrinsic Estimator for the model with age-, cohort- and year-dummies and $MORT$ as defined in regression equation (6) is presented; in column (2), $SR5$ is added

(regression equation (7)). Columns (3) to (6) then show the results for the dynamic panel model (regression equation (8)) with both $MORT$ and $SR5$, estimated by the difference-GMM-estimator due to Arellano and Bond (1991) and by the system-GMM-estimator due to Blundell and Bond (1998). In all the GMM-estimations, HCE_{t-1} and $SR5_t$ are regarded to be predetermined as they do not depend on the error term in period t. In (3) and (4) $MORT_t$ is assumed to be predetermined, too, while in (5) and (6) we allow for $MORT_t$ to be endogenous. To limit instrument proliferation, the number of instruments was reduced using the *collapse*-option of STATA's *xtabond2*-command, see Roodman (2006). Results with the full set of instruments are, however, very similar.

These six models are estimated with the variables HCE, $MORT$ and $SR5$ in levels – see columns (1) to (6) – and in first differences (columns (7) to (12) in the middle part of the table). Because not all unit root tests reject non-stationarity of the explanatory variables in first differences, we also present the six models in second differences (columns (13) to (18) in the bottom part of the table). However, for women the AR(2)-test is highly significant (with a p-value < 0.001 for the difference GMM-estimator, and 0.002 for the system GMM-estimator), which is a clear indicator that the model in second differences is misspecified for women, so we present results in second differences only for men.

We first observe that the coefficients of mortality are positive and highly significant for men. They suggest that expenditures for men in their last year of life are between 5 and 14 times as high as for the average sickness fund member. These estimates are roughly in line with findings from previous studies. E.g., Lubitz and Riley (1993) found that the 5 per cent decedents account for 25-30 per cent of total Medicare expenditures. The Lubitz-Riley results imply that decedents spend about 6 times as much as survivors. For women, the coefficients are positive, but usually smaller and not always significant.

Longevity, measured by the predicted value of the 5-year survival rate, always has a positive and significant impact on HCE, although the size of the coefficient varies according to the specification. A value of 12, which seems to be a lower bound (for men), suggests that an increase in the 5-year survival rate by 5 percentage points (which occurred for men over 70 and for women between 75 and 85 from 1997 to 2009) raises real daily per-capita HCE by roughly 10 per cent.

None of these results depends on whether the mortality rate is treated as predetermined or endogenous. If anything, the coefficient of mortality tends to be somewhat larger when mortality is treated as endogenous than otherwise.

We now turn to the results of the age, cohort and time dummies. We present graphs for the model in column (2) in Tables 4 and 5, with $MORT$ and $SR5$ as additional variables. In Figure 1, we observe that the age dummies show a familiar picture: a high value for newborns, then a decline up to age 3, followed by a relatively flat portion up to age 45 (with somewhat higher expenditures for women in child-bearing age), and then a steep rise until age 89. It is remarkable that this pattern remains even though both the mortality rate and the 5-year survival rate are held constant. Thus there seems to be an independent effect of age on HCE, in contrast to the findings of the previous literature.

The coefficients of the cohort dummies are declining except for the first and last few cohorts, which we observe only for a smaller number of years than the other cohorts, see Figure 2. The general pattern confirms the well-known fact that more recent cohorts are healthier at a given age and therefore need less medical care than older cohorts.

Table 4: Regression Results for Men, dependent variable: daily HCE in levels (upper part), in first differences (middle part), and in second differences (lower part)

	Men, Levels					
	(1)	(2)	(3)	(4)	(5)	(6)
GMM			Dif.	System	Dif.	Sys.
$MORT$ endog.					✓	✓
$MORT$	**68.26**	**58.30**	**70.17**	**30.01**	**75.56**	**37.25**
	(.000)	**(.000)**	**(.000)**	**(.000)**	**(.000)**	**(.000)**
$SR5$		**36.45**	**34.82**	**12.45**	**33.46**	**14.23**
		(.000)	**(.000)**	**(.000)**	**(.000)**	**(.000)**
HCE_{t-1}			**0.12**	**0.23**	**0.11**	**0.22**
			(.000)	**(.000)**	**(.008)**	**(.000)**
$AR(1)$			(.000)	(.000)	(.000)	(.000)
$AR(2)$			(.493)	(.605)	(.484)	(.627)
	Men, First Differences (Δ)					
	(7)	(8)	(9)	(10)	(11)	(12)
GMM			Dif.	Sys.	Dif.	Sys.
$MORT$ endog.					✓	✓
$\Delta MORT$	**60.86**	**56.22**	**56.54**	**55.03**	**83.78**	**77.18**
	(.000)	**(.000)**	**(.000)**	**(.000)**	**(.000)**	**(.000)**
$\Delta SR5$		**13.83**	**16.32**	**15.16**	**12.46**	**12.08**
		(.015)	**(.001)**	**(.022)**	**(.018)**	**(.049)**
ΔHCE_{t-1}			-0.02	-0.01	-0.03	-0.02
			(.583)	(.887)	(.310)	(.592)
$AR(1)$			(.000)	(.000)	(.000)	(.000)
$AR(2)$			(.401)	(.339)	(.312)	(.308)
	Men, Second Differences (Δ^2)					
	(13)	(14)	(15)	(16)	(17)	(18)
GMM			Dif.	Sys.	Dif.	Sys.
$MORT$ endog.					✓	✓
$\Delta^2 MORT$	**51.28**	**42.41**	**49.30**	**44.72**	**58.51**	**64.66**
	(.000)	**(.000)**	**(.000)**	**(.000)**	**(.000)**	**(.000)**
$\Delta^2 SR5$		**21.84**	**12.36**	**15.13**	**11.51**	**13.12**
		(.000)	**(.002)**	**(.002)**	**(.004)**	**(.006)**
$\Delta^2 HCE_{t-1}$			**-0.32**	**-0.25**	**-0.32**	**-0.23**
			(.000)	**(.000)**	**(.000)**	**(.000)**
$AR(1)$			(.000)	(.000)	(.000)	(.000)
$AR(2)$			(.602)	(.613)	(.970)	(.149)

p-values in parentheses; **bold figures**: significant at $\alpha = .05$.

Table 5: Regression Results for Women, dependent variable: daily HCE in levels (upper part), in first differences (middle part), and in second differences (lower part)

	Women, Levels					
	(1)	(2)	(3)	(4)	(5)	(6)
GMM			Dif.	System	Dif.	Sys.
$MORT$ endog.					✓	✓
$MORT$	27.22	**26.24**	15.20	**33.68**	15.25	**41.38**
	(.055)	**(.001)**	(.245)	**(.000)**	(.245)	**(.000)**
$SR5$		**42.69**	**29.01**	**8.99**	**28.99**	**10.66**
		(.000)	**(.000)**	**(.000)**	**(.000)**	**(.000)**
HCE_{t-1}			**0.28**	**0.31**	**0.28**	**0.28**
			(.000)	**(.000)**	**(.000)**	**(.000)**
$AR(1)$			(.000)	(.000)	(.000)	(.000)
$AR(2)$			(.973)	(.751)	(.973)	(.861)
	Women, First Differences (Δ)					
	(7)	(8)	(9)	(10)	(11)	(12)
GMM			Dif.	Sys.	Dif.	Sys.
$MORT$ endog.					✓	✓
$\Delta MORT$	**33.64**	20.38	3.24	4.99	8.46	7.90
	(.000)	(.053)	**(.703)**	**(.523)**	**(.431)**	**(.384)**
$\Delta SR5$		**15.77**	**19.97**	**18.22**	**19.43**	**18.01**
		(.000)	**(.000)**	**(.000)**	**(.000)**	**(.000)**
ΔHCE_{t-1}			**0.12**	**0.14**	**0.11**	**0.13**
			(.000)	**(.000)**	**(.000)**	**(.000)**
$AR(1)$			(.000)	(.000)	(.000)	(.000)
$AR(2)$			(.781)	(.689)	(.791)	(.697)

p-values in parentheses; **bold figures**: significant at $\alpha = .05$.

Figure 3 shows the positive time trend for HCE. It also shows the impact of a major health care reform that took effect in 2004. Depending on the model specification, the year dummies indicate an annual growth rate of 1.95 to 2.32 per cent for men and 1.02 to 1.62 per cent for women (see the last row of Table 6), which can be interpreted as the "pure time trend in real per-capita HCE", independent of demographic effects.

We conclude that the hypotheses stated in Section 3 are supported by the results for both sexes. Since both the mortality rate and longevity have a (mostly) significantly positive effect on HCE, the sign of the total effect of population ageing, which leads both to a decline in mortality and an increase in longevity, is unclear. Therefore, we have to use simulation methods to determine whether the total effect will be positive, given the demographic development predicted for Germany.

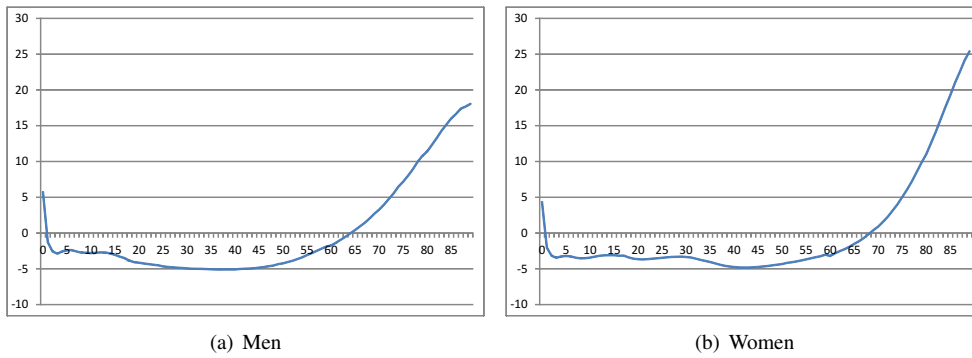Figure 1: Age dummy coefficients for regression equation (7)

(a) Men

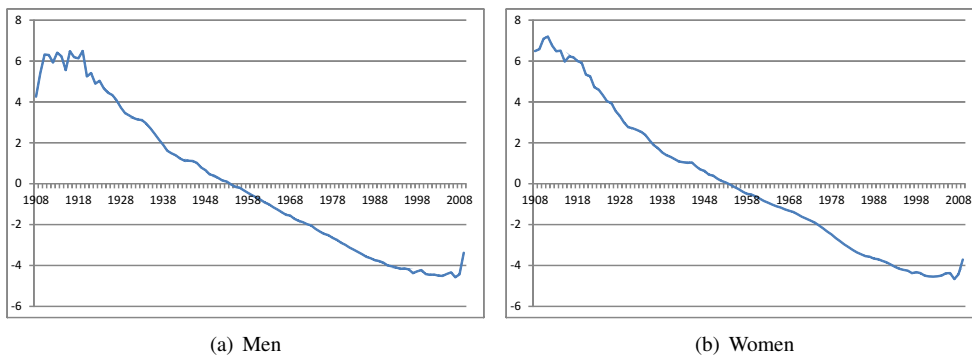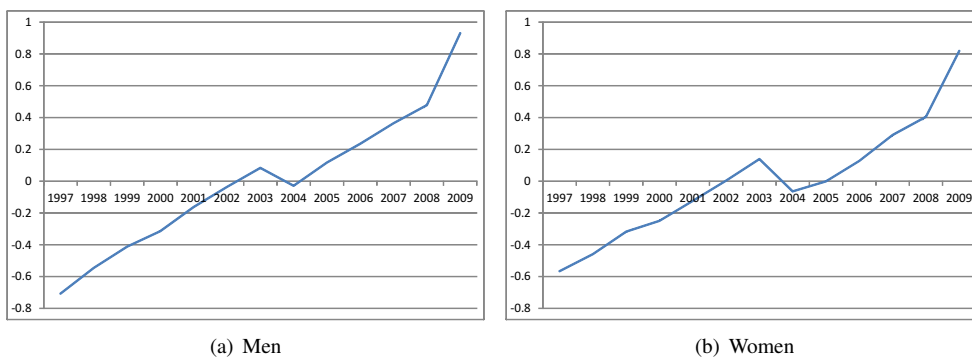

(b) Women

Figure 2: Cohort dummy coefficients for regression equation (7)

(a) Men

(b) Women

Figure 3: Year dummy coefficients for regression equation (7)

(a) Men

(b) Women

6 Estimating the Demographic Effect on Health Care Expenditures

In the following, we do not attempt to *forecast* the development of health care expenditures in Germany over the next decades. This would be a futile endeavour, because this depends to a great extent on political decisions. Instead, we are trying to measure the purely demographic impact on HCE by performing a counterfactual exercise in that we vary only the demographic factors, holding everything else constant at the 2009 level. For ease of interpretation, we divide the resulting values by the respective 2009 value of HCE, so that we can interpret the result as the relative increase of HCE due to demographic change.

We proceed in three steps. We first consider only the effect of the reduction of mortality rates (without its impact on the 5-year survival rates and the age distribution). To do so, we calculate the age profiles of HCE and per capita HCE that would result from changing only the mortality rates for all age groups to their values in 2020, 2030, 2040, 2050 and 2060, using the regression results of the models with only $MORT$ as an additional explanatory variable besides age, year and cohort. Columns (1), (7) and (13) of the upper part of Tables 6 and 7 show that the well-known red herring effect is present in our data as well: When the mortality rates decline in the way predicted for the next decades and everything else stays the same, the age profiles of HCE shift downwards because in each age bracket, fewer people are in their last year of life, so that per capita HCE decrease. However, the overall impact is rather modest: With the mortality rates of 2060, expenditures in 2009 for men would have been lower by at most 7.1 per cent, those for women by 3.2 per cent. Note that the calculations in this first step (columns (1), (7) and (13)) serve only as a benchmark for comparison because considering the change in mortality and ignoring the concomitant increase in survival rates of the elderly is inconsistent.

In the second step, we take into account that with falling mortality the 5-year survival rates must rise, which by itself would raise HCE. We therefore calculate the age profiles of HCE and per capita HCE that would result from changing both the mortality rates and the 5-year survival rates to their values in 2020, 2030, ... 2060, see columns (2)-(6), (8)-(12), and (14)-(18) in the upper part of Table 6 and columns (2)-(6) and (8)-(12) in the upper part of Table 7. For men, the total change in HCE resulting from this variation lies between minus 2.3 per cent (column (11) of Table 6) and plus 12.6 per cent (column (2)). For women, the total change is always positive and lies between 1 and 17 per cent (columns (4) and (2) of Table 7, respectively).

Thus we see that, depending on the estimation method used, the decline in HCE due to lower mortality rates is either considerably mitigated or more than compensated by considering the concomitant increase in the 5-year survival rates of older population groups. We emphasize again that these results do not at all depend on how the multicollinearity problem is solved.

In the third step, we also set the age distribution to their levels in 2020 through 2060. These results must be interpreted with caution because when we make use of the age dummy coefficients, we also have to decide how to treat the coefficients of the cohort dummies. However, there is no natural way to extrapolate the cohort effects because it is not known how healthy or unhealthy future cohorts will be. To make matters worse, there is no monotone trend in the cohort coefficients which could be easily extrapolated (see Figure 2). We therefore did not use any predicted (extrapolated) values for the cohorts but left them at their 2009 values, but this is not much more than the application of the Principle of Insufficient Reason. The results of this exercise can be found in the lower part of Tables 6 and 7. The numbers show that with the 2060 age composition (along with the 2060 mortality and survival rates), health care expenditures in

Table 6: Relative values of per capita HCE when mortality rates and survival rates (and the age distribution) are set to their future values

Men

	(1) level	(2) level	(3) level Dif.	(4) level Sys.	(5) level Dif.	(6) level Sys.	(7) Δ	(8) Δ	(9) Δ Dif.	(10) Δ Sys.	(11) Δ Dif.	(12) Δ Sys.	(13) Δ^2	(14) Δ^2	(15) Δ^2 Dif.	(16) Δ^2 Sys.	(17) Δ^2 Dif.	(18) Δ^2 Sys.
GMM MORT endog.					✓	✓					✓	✓					✓	✓
$MORT$	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
$SR5$		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
HCE_{t-1}			✓	✓	✓	✓			✓	✓	✓	✓			✓	✓	✓	✓
Age distribution not adjusted																		
2009	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2020	0.975	1.033	1.026	1.007	1.022	1.007	0.978	1.000	1.004	1.002	0.988	0.990	0.981	1.017	1.000	1.006	0.996	0.996
2030	0.960	1.061	1.050	1.015	1.044	1.016	0.964	1.003	1.010	1.008	0.984	0.987	0.970	1.033	1.004	1.014	0.996	0.997
2040	0.948	1.085	1.071	1.022	1.062	1.022	0.954	1.007	1.015	1.012	0.981	0.984	0.961	1.046	1.007	1.020	0.997	0.998
2050	0.938	1.107	1.089	1.027	1.078	1.029	0.944	1.010	1.020	1.016	0.978	0.983	0.953	1.057	1.009	1.026	0.997	0.999
2060	0.929	1.126	1.105	1.033	1.093	1.034	0.936	1.012	1.025	1.020	0.977	0.981	0.946	1.068	1.012	1.031	0.998	1.000
Age distribution adjusted to 2060																		
2009	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2020	1.078	1.152	1.144	1.119	1.139	1.119	1.081	1.110	1.114	1.113	1.095	1.097	1.085	1.131	1.110	1.117	1.104	1.104
2030	1.137	1.286	1.270	1.217	1.260	1.218	1.143	1.201	1.210	1.207	1.173	1.177	1.151	1.243	1.201	1.215	1.190	1.191
2040	1.184	1.419	1.393	1.310	1.378	1.311	1.193	1.284	1.299	1.294	1.240	1.246	1.206	1.351	1.284	1.307	1.267	1.269
2050	1.190	1.505	1.471	1.359	1.450	1.361	1.203	1.324	1.344	1.337	1.264	1.273	1.220	1.414	1.325	1.355	1.301	1.304
2060	1.184	1.554	1.514	1.380	1.491	1.383	1.199	1.341	1.364	1.356	1.273	1.282	1.218	1.446	1.341	1.376	1.314	1.317
%-growth rate demographic	0.33	0.87	0.82	0.63	0.79	0.64	0.36	0.58	0.61	0.60	0.47	0.49	0.39	0.73	0.58	0.63	0.54	0.54
%-growth rate time trend	1.95	2.32	2.28	2.13	2.26	2.14	1.96	2.10	2.13	2.12	2.05	2.06	1.98	2.20	2.10	2.13	2.08	2.08

16

Table 7: Relative values of per capita HCE when mortality rates and survival rates (and the age distribution) are set to their future values

Women

	(1) level	(2) level	(3) level Dif.	(4) level Sys.	(5) level Dif.	(6) level Sys.	(7) Δ	(8) Δ	(9) Δ Dif.	(10) Δ Sys.	(11) Δ Dif.	(12) Δ Sys.
GMM												
$MORT$ endog.	✓											
$MORT$		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
$SR5$		✓	✓	✓	✓	✓		✓	✓	✓	✓	✓
HCE_{t-1}			✓	✓	✓	✓			✓	✓	✓	✓
Age distribution not adjusted												
2009	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2020	0.991	1.048	1.033	1.001	1.033	1.001	0.989	1.014	1.025	1.023	1.023	1.021
2030	0.986	1.084	1.059	1.003	1.059	1.003	0.983	1.026	1.044	1.039	1.040	1.037
2040	0.982	1.116	1.081	1.006	1.081	1.006	0.977	1.036	1.061	1.054	1.056	1.051
2050	0.978	1.145	1.101	1.008	1.101	1.008	0.972	1.045	1.076	1.067	1.069	1.064
2060	0.974	1.170	1.119	1.010	1.118	1.010	0.968	1.053	1.089	1.079	1.081	1.075
Age distribution adjusted to 2060												
2009	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2020	1.058	1.126	1.108	1.070	1.108	1.070	1.056	1.085	1.097	1.094	1.095	1.093
2030	1.107	1.237	1.202	1.130	1.202	1.131	1.103	1.159	1.181	1.175	1.176	1.172
2040	1.160	1.375	1.317	1.199	1.317	1.200	1.154	1.246	1.284	1.273	1.276	1.269
2050	1.191	1.487	1.408	1.245	1.408	1.246	1.183	1.310	1.361	1.347	1.351	1.342
2060	1.192	1.532	1.440	1.254	1.440	1.256	1.182	1.327	1.385	1.369	1.374	1.363
%-growth rate demographic	0.34	0.84	0.72	0.44	0.72	0.45	0.33	0.56	0.64	0.62	0.62	0.61
%-growth rate time trend	1.02	1.62	1.46	1.14	1.46	1.14	1.01	1.26	1.36	1.33	1.34	1.32

17

2009 would have been between 27 and 55 per cent higher for men and between 25 and 53 per cent higher for women, an effect that is considerably higher than the impact of mortality and survival rates alone. The second line from the bottom in each of the Tables 6 and 7 contains the results of converting the respective increases into annual growth rates, which can be interpreted as "growth in HCE due to demographic change". Considering both the changes in mortality and in 5-year survival rates, these numbers lie between .47 and .87 for men and between .44 and .84 per cent for women.

In the last line of Tables 6 and 7 we present the pure time trend in real per-capita HCE, independent of demographic effects, calculated from the coefficients for the year dummies. It can be assumed that this trend is to a great extent due to medical progress. The annual growth rates lie slightly above 2 per cent for men and between 1.1 and 1.6 per cent for women and are thus considerably larger than the purely demographic effect estimated above. If these two effects are added up, the resulting growth rates lie between 2.5 and 3 per cent for men and between 1.5 and 2.5 per cent for women, which is somewhat higher than common forecasts of the growth rate of per capita income in the ageing German population. Thus they suggest that demographic change and technical progress combined may after all present problems for the financing of health care in Germany.

7   Conclusions and Caveats

In this paper, we have used a pseudo-panel of HCE data for Germany to demonstrate that per-capita health care expenditures are significantly influenced by the age composition of the population, by mortality rates and by the development of longevity, as measured by the age-specific 5-year survival rates. We believe that the last effect, which is quite substantial, mirrors the medical profession's willingness to perform expensive or risky treatments on elderly patients if the patients can be expected to live long enough to enjoy the effects of the treatment.

The results of the simulations based on the regression coefficients show that if past trends continue, per-capita health care expenditures would rise by between 1.5 and 2 per cent per year even without demographic change. Moreover, while we can confirm that simulations on the basis of the population age structure alone are misleading, the same applies when only age-specific mortality rates are added. The effect of rising longevity can not be ignored, either. One way to take it into account is to include a measure of age-specific 5-year survival rates. In sum, the (negative) effect of falling mortality rates on health care expenditures is at least compensated and probably over-compensated by the (positive) effect of increasing 5-year survival rates. Adding the effect of a changing age composition of the population, the total effect of demographic change on health expenditures is estimated to amount to an annual growth rate between .4 and .9 per cent, depending on which estimator is used.

The type of data employed for this study has important advantages, but also certain drawbacks. To our knowledge, this is the first attempt to quantify the effect of rising longevity on the development of health care expenditures over time. However, since we used age and sex group averages instead of individual expenditure data, the well-known end-of-life effect on HCE expenditures is accounted for only in an indirect form: by estimating the impact of the mortality *rate* within a population group on *average* expenditures.

It can further be argued that mortality and survival rates themselves are influenced by HCE and therefore endogenous. With respect to $SR5$, the endogeneity does not occur as we used its predicted value instead of $SR5$ itself. For $MORT$, possible endogeneity is accounted for in two of the four dynamic panel models (estimated by GMM), which had basically no effect on

the regression results. This seems reasonable as one may argue that, unlike in individual data, for group averages the causal effect of HCE on mortality should not be too strong. It does not seem likely that the correlation of the variation in HCE and $MORT$ is caused primarily by the fact that tight rationing against a particular age-sex group as a whole in a certain year by all physicians led to a higher mortality rate, but rather by a higher mortality rate of an age-sex group causing higher expenditures.

We sum up by stating the main purpose of this paper, namely to examine whether ageing – i.e. an increase of longevity alongside a fall in mortality rates – as such will increase health expenditures, and the answer to this question is a clear "yes". Moreover, independent of the specification, the 5-year survival rate always has a positive and sizeable impact on health care expenditures so that for Germany a Eubie Blake effect indeed exists.

Appendix

The following Tables 8 to 10 provide the unit root tests for the variables HCE, $MORT$ and $SR5$.

Table 8: Unit root tests, dep. variable: daily HCE

Men

time trend	HCE yes		ΔHCE no		ΔHCE yes		Δ²HCE no		Δ²HCE yes	
	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value
Harris-Tsavalis $\bar{t}$	0.316	0.000	0.202	0.000	0.486	0.951	-0.124	0.000	-0.093	0.000
Im-Pesaran-Shin $W_{\bar{t}}$	-2.582	<0.01	-3.507	<0.01	-3.679	<0.01	-4.844	<0.01	-4.756	<0.01
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 1)	-1.356	0.088	-16.022	0.000	-10.518	0.000	-27.355	0.000	-18.079	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 2)	-1.745	0.041	-15.666	0.000	-9.788	0.000	-25.103	0.000	-17.280	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 3)	-1.339	0.090	-11.258	0.000	-5.678	0.000	-17.004	0.000	-9.909	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 4)	-44.188	0.000	-38.012	0.000	-2.502	0.006	-12.344	0.000	-9.652	0.000

Women

time trend	HCE yes		ΔHCE no		ΔHCE yes		Δ²HCE no		Δ²HCE yes	
	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value
Harris-Tsavalis $\bar{t}$	0.355	0.001	0.188	0.000	0.404	0.308	-0.159	0.000	-0.125	0.000
Im-Pesaran-Shin $W_{\bar{t}}$	-2.958	<0.01	-3.657	<0.01	-4.027	<0.01	-5.192	<0.01	-5.018	<0.01
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 1)	-5.460	0.000	-16.221	0.000	-14.615	0.000	-29.605	0.000	-19.986	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 2)	-7.541	0.000	-15.188	0.000	-10.013	0.000	-21.788	0.000	-13.408	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 3)	-3.702	0.000	-11.984	0.000	-2.270	0.012	-14.892	0.000	-66.537	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 4)	-36.543	0.000	-45.483	0.000	-1.461	0.072	-12.543	0.000	-10.228	0.000

IPS-test with lags contains optimal number of lags (up to the maximal number of lags given in parenthesis) according to BIC.

Table 9: Unit root tests, dep. variable: $MORT$

Men

time trend	$MORT$ yes		$\Delta MORT$ no		$\Delta MORT$ yes		$\Delta^2 MORT$ no		$\Delta^2 MORT$ yes	
	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value
Harris-Tsavalis $\bar{t}$	1.062	1.000	0.378	0.000	0.987	1.000	-0.782	0.000	-0.702	0.000
Im-Pesaran-Shin $W_{\bar{t}}$	0.477	>0.10	-1.538	>0.10	-3.924	<0.01	-6.493	<0.01	-6.516	<0.01
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 1)	26.843	1.000	9.075	1.000	-15.452	0.000	-42.604	0.000	-31.978	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 2)	28.353	1.000	12.626	1.000	-13.437	0.000	-39.433	0.000	-30.945	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 3)	28.149	1.000	12.995	1.000	-10.137	0.000	-35.249	0.000	-25.113	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 4)	25.275	1.000	17.103	1.000	-9.912	0.000	-34.699	0.000	-4.021	0.000

Women

time trend	$MORT$ yes		$\Delta MORT$ no		$\Delta MORT$ yes		$\Delta^2 MORT$ no		$\Delta^2 MORT$ yes	
	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value
Harris-Tsavalis $\bar{t}$	1.144	1.000	0.793	0.934	1.072	1.000	-0.600	0.000	-0.390	0.000
Im-Pesaran-Shin $W_{\bar{t}}$	1.327	>0.10	-0.740	>0.10	-2.983	<0.01	-5.045	<0.01	-5.273	<0.01
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 1)	31.784	1.000	13.085	1.000	-6.146	0.000	-26.354	0.000	-18.345	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 2)	32.618	1.000	19.735	1.000	-2.449	0.007	-22.604	0.000	-16.391	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 3)	31.344	1.000	19.887	1.000	-2.364	0.009	-21.566	0.000	-14.311	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 4)	26.890	1.000	20.684	1.000	-1.079	0.140	-18.747	0.000	-3.986	0.000

IPS-test with lags contains optimal number of lags (up to the maximal number of lags given in parenthesis) according to BIC.

Table 10: Unit root tests, dep. variable: daily $SR5$

Men

time trend	$SR5$ yes		$\Delta SR5$ no		$\Delta SR5$ yes		$\Delta^2 SR5$ no		$\Delta^2 SR5$ yes	
	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value
Harris-Tsavalis $\bar{t}$	0.943	1.000	0.617	0.000	1.0488	1.000	-0.480	0.000	-0.395	0.000
Im-Pesaran-Shin $W_{\bar{t}}$	-0.205	>0.10	-0.899	>0.10	-3.170	<0.01	-4.052	<0.01	-3.856	<0.01
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 1)	17.368	1.000	5.993	1.000	-8.429	0.000	-19.511	0.000	-10.829	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 2)	17.077	1.000	6.240	1.000	-7.698	0.000	-18.645	0.000	-9.074	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 3)	15.217	1.000	5.904	1.000	-5.477	0.000	-13.483	0.000	-10.088	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 4)	0.987	0.838	1.915	0.972	-5.940	0.000	-13.653	0.000	-9.422	0.000

Women

time trend	$SR5$ yes		$\Delta SR5$ no		$\Delta SR5$ yes		$\Delta^2 SR5$ no		$\Delta^2 SR5$ yes	
	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value	Stat.	p-value
Harris-Tsavalis $\bar{t}$	0.981	1.000	0.897	1.000	1.086	1.000	-0.519	0.000	-0.264	0.000
Im-Pesaran-Shin $W_{\bar{t}}$	0.211	>0.10	-0.187	>0.10	-3.065	<0.01	-3.704	<0.01	-3.579	<0.01
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 1)	20.869	1.000	11.922	1.000	-7.553	0.000	-16.465	0.000	-8.743	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 2)	20.519	1.000	11.333	1.000	-7.368	0.000	-16.040	0.000	-7.970	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 3)	18.807	1.000	11.074	1.000	-8.300	0.000	-11.062	0.000	-3.903	0.000
Im-Pesaran-Shin $W_{\bar{t}}$ (lag up to 4)	11.269	1.000	9.930	1.000	-8.413	0.000	-11.404	0.000	-8.151	0.000

IPS-test with lags contains optimal number of lags (up to the maximal number of lags given in parenthesis) according to BIC.

References

ARELLANO, M., AND S. BOND (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, 58, 277–287.

BECH, M., T. CHRISTIANSEN, E. KHOMAN, J. LAURIDSEN, AND M. WEALE (2011): "Ageing and health care expenditure in EU-15," *European Journal of Health Economics*, 12, 469–478.

BLUNDELL, R., AND S. BOND (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87, 115–143.

BREYER, F., AND S. HUPFELD (2009): "Fairness of Public Pensions and Old-Age Poverty," *FinanzArchiv*, 65, 358–380.

CALLAHAN, D. (1987): *Setting Limits: Medical Goals in an Aging Society*. New York.

DANIELS, N. (1985): *Just Health Care*. Camebridge.

DAVIDSON, R., AND J. G. MACKINNON (1993): *Estimation and Inference in Econometrics*. New York.

DEATON, A. (1985): "Panel Data from Time Series of Cross-Sections," *Journal of Econometrics*, 30, 109–126.

FANG, H., M. KEANE, A. KHWAJA, M. SALM, AND D. SILVERMAN (2007): "Testing the Mechanisms of Structural Models: The Case of the Mickey Mantle Effect," *American Economic Review: Papers and Proceedings*, 97(2), 53–59.

FUCHS, V. (1984): "Though much is taken: reflections on aging, health and medical care," *Milbank Memorial Fund Quarterly/Health and Society*, 61, 143–1.

HAGIST, C., N. KLUSEN, A. PLATE, AND B. RAFFELHÜSCHEN (2005): *Social Health Insurance – the Major Driver of Unsustainable Fiscal Policy?*, CESifo Working Paper; No. 1574.

HAGIST, C., AND L. KOTLIKOFF (2005): *Who's Going Broke? Comparing Healthcare Costs in ten OECD Countries*, NBER Working Paper; w11833.

HARRIS, R., AND E. TZAVALIS (1999): "Inference for unit roots in dynamic panels where the time dimension is fixed," *Journal of Econometrics*, 91, 201–226.

HUMAN MORTALITY DATABASE (2011): University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org or www.humanmortality.de (data downloaded on May 31, 2011).

IM, K., M. PESARAN, AND Y. SHIN (2003): "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 115, 53–74.

INKMANN, J., S. KLOTZ, AND W. POHLMEIER (1998): "Permanente Narben oder temporäre Blessuren? Eine Studie über die langfristigen Folgen eines missglückten Einstiegs in das Berufsleben auf der Grundlage von Pseudo-Panel-Daten," in *Qualifikation, Weiterbildung und Arbeitsmarkterfolg*, ed. by F. Pfeiffer, and W. Pohlmeier, pp. 223–255. Nomos, Baden-Baden.

KARLSSON, M., AND F. KLOHN (2011): "Some notes on how to catch a red herring. Ageing, time-to-death & care costs for older people in Sweden," Darmstadt discussion papers in economics, Darmstadt Technical University, Department of Business Administration, Economics and Law, Institute of Economics (VWL).

LUBITZ, J., J. BEEBE, AND C. BAKER (1995): "Longevity and Medicare expenditure," *New England Journal of Medicine*, 332, 999–1003.

LUBITZ, J., AND G. F. RILEY (1993): "Trends in Medicare Payments in the Last Year of Life," *New England Journal of Medicine*, 328, 1092–1096.

MILLER, T. (2001): "Increasing Longevity and Medicare Expenditures," *Demography*, 38, 215–226.

ROODMAN, D. (2006): "How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata," *Center for Global Development Working Paper*, 103.

SESHAMANI, M., AND A. GRAY (2004): "Ageing and health-care expenditure: the red herring argument revisited," *Health Economics*, 13, 303–314.

SHANG, B., AND D. GOLDMAN (2008): "Does Age or Life Expectancy Better Predict Health Care Expenditures?," *Health Economics*, 17, 487–501.

STATISTISCHES BUNDESAMT (2009): *12. koordinierte Bevölkerungsvorausberechnung*. Wiesbaden.

STEARNS, S., AND E. NORTON (2004): "Time to include time to death? The future of health care expenditure predictions," *Health Economics*, 13, 315–327.

STRECH, D., M. SYNOFZIK, AND G. MARCKMANN (2008): "How Physicians Allocate Scarce Resources at the Bedside: A Systematic Review of Qualitative Studies," *Journal of Medicine and Philosophy*, 33, 80–99.

TU, Y.-K., N. KRÄMER, AND W.-C. LEE (2012): "Addressing the Identification Problem in Age-period-cohort Analysis," *Epidemiology*, 23(3), 1–11.

VERBEEK, M., AND T. NIJMAN (1992): "Can Cohort Data be Treated as Genuine Panel Data?," *Empirical Economics*, 17, 9–23.

VON GAUDECKER, H.-M., AND R. SCHOLZ (2007): "Differential Mortality by Lifetime Earnings in Germany," *Empirical Economics*, 17, 83–108.

WERBLOW, A., S. FELDER, AND P. ZWEIFEL (2007): "Population ageing and health care expenditure: a school of red herrings?," *Health Economics*, 146, 1109–1127.

YANG, Y., S. SCHULHOFER-WOHL, W. FU, AND K. LAND (2008): "The Intrinsic Estimator for Age-Period-Cohort Analysis: What it is and How to Use It," *American Journal of Sociology*, 113(6), 1697–1736.

ZWEIFEL, P., S. FELDER, AND M. MEIER (1999): "Ageing of population and health care expenditure: a red herring?," *Health Economics*, 8, 485–496.

ZWEIFEL, P., S. FELDER, AND A. WERBLOW (2004): "Population ageing and health care expenditure: New Evidence on the Red Herring," *Geneva Papers on Risk and Insurance: Issues and Practice. Special Issue on Health Insurance*, 29(4), 653–667.

ZWEIFEL, P., L. STEINMANN, AND P. EUGSTER (2005): "The Sisyphus Syndrome in Health Revisited," *International Journal of Health Care Finance and Economics*, 5, 127–145.