The Geneva Papers on Risk and Insurance Vol. 29 No. 4 (October 2004) 652–666

Population Ageing and Health Care Expenditure: New Evidence on the "Red Herring"

by Peter Zweifel, Stefan Felder and Andreas Werblow*

The observation that average health care expenditure rises with age generally leads experts and laymen alike to conclude that population ageing is the main driver of health care costs. In recently published studies we challenged this view (Zweifel *et al.,* 1999; Felder *et al.,* 2000). Analysing health care expenditure of deceased persons, we showed that age is insignificant if proximity to death is controlled for. Thus, we argued that population ageing *per se* will not have a significant impact on future health care expenditure. Several authors (Salas and Raftery, 2001; Dow and Norton, 2002; Seshamani and Gray, 2004a) disputed the robustness of these findings, pointing to potential weaknesses in the econometric methodology. This paper revisits the debate and provides new empirical evidence, taking into account the methodological concerns that have been raised. We also include surviving individuals to test for the possibility that the relative importance of proximity to death and age differs between the deceased and survivors. The results vindicate our earlier findings of no significant age effect on health care expenditure of the deceased. However, with respect to the survivors, we find that age may matter. Still, a naïve estimation that does not control for proximity to death will grossly overestimate the effect of population ageing on aggregate health care expenditure. Following Stearns and Norton (2004), we conclude that "it is time for time to death" in projections of future health care costs.

1. Introduction

In the debate over ever-increasing health care costs, a routine argument states that ageing of the population is a crucial driver of health care expenditure (HCE). Zweifel, Felder and Meier (1999) (hereafter "ZFM") argued that this claim is a red herring. Their dissent was based on the analysis of health care expenditure of deceased persons in their last years of life. The number of quarters remaining until death was significant while the age of the persons was not. These findings have been vindicated by a number of subsequent studies (e.g. O'Neill *et al.*, 2000). Several authors, however, have raised methodological concerns about the ZFM model (Dow and Norton, 2002; Salas and Raftery, 2001; Seshamani and Gray, 2004a). Their main arguments refer to multicollinearity between and endogeneity of the explanatory variables.

If present, multicollinearity is indeed a problem given that the null hypothesis states age to be insignificant for a person's HCE. Collinearity between two or more explanatory variables increases the variance of estimated regression coefficients. Thus, the higher the degree of multicollinearity, the higher the probability of type II error (rejecting the alternative hypothesis of a significant effect when in fact it is true). While this effect applies to all regressors, it may be particularly marked with regard to the age-related ones.

* Peter Zweifel is from the Socioeconomic Institute, University of Zurich, Hottingerstr. 10, CH-8032 Zurich, email: pzweifel@soi.unizh.ch; Stephan Felder and Andreas Werblow are from the Institute of Social Medicine and Health Economics (ISMHE), Otto-von-Guericke-University Magdeburg, Leipziger Str. 44, D-39120 Magdeburg. Corresponding author email: stefan.felder@ismhe.de

In the present context, there are two sources of multicollinearity:

1. HCE are censored, zero-inflated and roughly log-normally distributed data. An option to deal with these peculiarities is the Heckman model, in which the probability of a positive HCE is estimated in the first step, followed by an ordinary least squares (OLS) estimation of logged HCE which includes the inverse Mill's ratio λ as an additional regressor. Dow and Norton (2002) as well as Salas and Raftery (2001) pointed to the possibility of a sizable correlation between λ and the other exogenous variables, in particular age.
2. Collinearity between age and age squared. This is specifically severe when the age range is rather small. In the original ZFM study, age-related variables are insignificant in the sub-sample of 65+ aged persons, while being significant in the overall sample. This difference might be due to a large degree of collinearity between the age variables (but also between λ and the age variables).

The second methodological critique of the ZFM approach concerns the potential endogeneity of several regressors. The regression included quarter dummies representing time to death. Specifically, Salas and Raftery (2001) argued that time to death ("TTD") might be influenced by current and previous HCE. In particular, if HCE should be autocorrelated over time, these dummies contain an error that is also part of the dependent variable, HCE. This endogeneity causes OLS estimates to yield biased and inconsistent coefficients.

In the present paper, multicollinearity is at least mitigated by employing a two-part model in addition to the Heckman model. The two-part model separates the selection part (estimation of the probability of a positive HCE) from the equation that explains the level of HCE. Thus, the correlation between the selection term λ and the age variables as a source of multicollinearity is eliminated.

We tackle the endogeneity problem by implementing a different approach from ZFM. Instead of using a panel data set (which gave rise to the set of quarter dummies each of which may be the cause of an endogeneity problem), we study HCE in one year and introduce time to death measured from that year as a single explanatory variable. In this way, HCE spent during the current year would have to influence TTD by increasing the probability of survival at the far end of the path, which is likely only if that far end is close, i.e. in the few cases of life-saving interventions. Also, the regression equation is modified in a way that has similarities with instrumental variable (IV) estimation. An additional concern has always been that the relative importance of TTD and age may differ between decedents and survivors. To test for this possibility, the analysis is repeated using a data set that comprises both deceased and surviving individuals. For the latter, proximity to death by necessity can be measured with substantial error only. This should benefit the age-related explanatory variables, thus resulting in an upper bound estimate of their impact on HCE.

The remainder of the paper is organized as follows: section 2 presents the sets of data, including the descriptive statistics of the main variables. Section 3 recapitulates our earlier approach (ZFM, 1999) and addresses multicollinearity and endogeneity of the regressors. Section 4 presents the results of the extended model that includes both survivors and the deceased. Section 5 summarizes.

2. Data

We employ two different sets of data from a Swiss sickness fund (see Table 1). The first set is used to recapitulate the original investigation and to address issues of multicollinearity and endogeneity. It contains monthly data on HCE in the last two years prior to death of

1,095 persons who died in 1999. Since the month of death but not the precise date was available, we deleted HCE in the last month of life and started observation in the month previous to the last month of life. This yields 23 observations of monthly HCE.

The second set of data includes 91,327 individuals enrolled in the sickness fund by 1 January 1997, and still alive and enrolled at the end of 30 June 2003 (87,543 individuals), or who died in the meantime (3,784 individuals). Their (annual) HCE information relates to the year 1999 and comes from all the inhabitants of two cantons, Geneva and Zurich, that were enrolled in this sickness fund. Given the large size of the sample, the relative number as well as the age structure of survivors and deceased appear to be fairly representative for the two regions.

Table 1:
Data sets used

Data set	Number of survivors	Number of deceased	HCE data	Frequency
1	None	1,095 individuals who died in 1999	Last two years prior to date of death	Monthly
2	87,543 individuals alive at 30 June 2003	3,784 individuals who died between 1 Jan. 2000 and 30 June 2003	1999	Yearly

Table 2 characterizes the individuals in the two sets of data. In the first sample, the average monthly HCE in the last 20 months of life is 1,284 Swiss Francs (SFr, some U.S.$ 850 at 2003 exchange rates). Of the 21,900 monthly HCE observations, 37 per cent (8,098) were zero. Average HCE of the deceased in 1999 was at SFr 18,258 in their last year of life. As to the second sample, HCE in 1999 of those who died between 1 January 2000 and 30 June 2003 was SFr 12,979, a figure 29 per cent below the average HCE in the first sample. In keeping with the maintained hypothesis, this difference may be attributed to the fact that for the majority of the deceased in sample two, the remaining life time in 1999 exceeds one year. Thus, for most of them HCE in 1999 does not include the high expenditure during their last months of life. Among the survivors, average HCE in 1999 was only SFr 2,390, again likely because time of death was still far away for most of them. The average age at death is 77 years, while the average age of the survivors is 43 years. The share of men is 42 per cent, with no difference between survivors and deceased. Zurich's share among the deceased is lower than its share among the survivors, indicating a younger population in Zurich.

The remaining variables in Table 2 characterize the insurance contracts of the individuals. Accident insurance can be bought in combination with health insurance, an option that is often chosen by the elderly. Individuals in the labour force usually obtain accident insurance through their employer. This explains why the share of individuals combining health and accident insurance within the same policy is lower among the survivors. Roughly one-third of the insured opted for hospital supplementary insurance, providing for amenities. Between 82 per cent and 94 per cent chose at least one further supplement to their health insurance policy, the higher share relating to the survivors. Swiss

social health insurance law allows individuals to choose higher deductibles. Among the deceased, one-fifth opted for high-deductible contracts, compared to one-third among the survivors.

Table 2:
Descriptive statistics of samples

	Deceased					
	1999 (n = 1,095)		2000–2003 (n = 3,784)		Survivors (n = 87,524)	
Variable	**Mean**	**SE**	**Mean**	**SE**	**Mean**	**SE**
HCE (in 1999)	18,258[a]	15,299[a]	12,979	14,574	2,390	5,100
Age	76.93	15.96	76.30	14.53	42.92	22.12
Time to death in months	12[a]	0[a]	20.75	11.97	>42	0
Share of men	0.42	0.49	0.41	0.49	0.43	0.50
Share of individuals from Zurich	0.68	0.47	0.69	0.46	0.74	0.44
Share of patients						
with accident insurance	0.95	0.23	0.93	0.25	0.72	0.45
with suppl. hospital insurance	0.31	0.46	0.32	0.47	0.34	0.47
with other supplements	0.82	0.39	0.84	0.36	0.94	0.24
with higher deductibles	0.20	0.40	0.22	0.41	0.33	0.47

[a] The figure relates to the HCE in the last year of life.

3. Detecting multicollinearity and endogeneity

We first replicate ZFM (1999) by again estimating a Heckman sample selection model (Heckman, 1979), which in a first step represents a selection mechanism that is responsible for some individuals having zero and other positive HCE. In this first step, we accordingly ran a probit model for estimating the likelihood of positive HCE and calculated the inverse Mill's ratio λ (which roughly indicates the extent to which an individual is over- or under-represented in the subsample with positive HCE). In the second step, we examined the level of HCE conditional on a positive level of HCE. The same regressors of the first part of the model were used in an OLS regression, and λ was inserted as an additional regressor to control for any potential biases caused by excluding zero HCE observations. In contradistinction to ZFM (1999), we used monthly rather than quarterly data, counting months backwards from the date of death.[1] Therefore, HCE for individual *i* in month *m* prior to death is related to the explanatory variables in the following probit and OLS models:

[1] A monthly specification of the date of death is preferable since it can follow the structure of HCE previous to death more closely than a quarterly specification could do.

$$\Pr(\mathrm{HCE_{im}} > 0) = \beta_0 + \beta_1 A_i + \beta_2 \frac{A_i^2}{1000} + \beta_3 S_i + \beta_4 (A_i \cdot S_i) + \beta_5 REG_i + \beta_6 HOSP_i + \sum_{j=1}^{19} \gamma_j M_{ij} + \varepsilon_{im}, \quad (1)$$

$$\ln HCE_{im} = \beta_0 + \beta_1 A_i + \beta_2 \frac{A_i^2}{1000} + \beta_3 S_i + \beta_4 (A_i \cdot S_i) + \beta_5 REG_i + \beta_6 HOSP_i + \beta_7 \lambda + \sum_{j=1}^{19} \gamma_j M_{ij} + \varphi_{im}, \quad (2)$$

where $A =$ patient age, $S =$ dummy variable for sex (with male $= 1$, female $= 0$), $REG =$ dummy variable for Cantons (Zurich $= 1$, Geneva $= 0$), $HOSP =$ dummy variable for supplementary hospital insurance, $M =$ dummy variables for up to 20 months prior to death.

We find a significant effect of proximity to death on the likelihood of incurring positive HCE in a given month (see Table 3). Months one through 12 prior to death have a significantly higher likelihood of positive HCE than month 20 from death. The age coefficients are both significant, with age imparting a negative but age-squared imparting a positive effect on the likelihood of a positive HCE.

In the equation for ln HCE, the OLS coefficients for age are insignificant, confirming the ZFM finding of no age effect on HCE if proximity to death is controlled for. Conversely, the month dummies are significant for months one to 12 prior to death, with a substantial upward trend towards the date of death, again in accordance with ZFM. However, we find no significant effect of λ on HCE.

Based on this replication of ZFM (1999), the two main criticisms of their approach are now addressed. The first is the issue of multicollinearity between the inverse Mill's ratio λ and exogenous regressors. Multicollinearity means near perfect linear dependence between explanatory variables which causes an inability to distinguish between them (indicated by high standard errors of coefficients and hence lack of significance). With λ artificially insignificant, the second step cannot be set apart from the first one. The use of the Heckman sample selection model for dealing with zero observations of HCE has been criticized for this reason, most forcefully by Dow and Norton (2002). One can demonstrate that if the same set of regressors is used in the two parts of the Heckman model, then the second part is only identified if the inverse Mill's ratio is a non-linear function of the regressors (see Vella, 1998, as well as Puhani, 2000). For identification in the case of linearity, one needs an explanatory variable that is significant in the selection, but not in the second part of the model.[2]

Indeed, an OLS regression of the inverse Mill's ratio λ on the explanatory variables results in an R^2 of 0.9897, suggesting almost perfect linearity.[3] This means that collinearity among the regressors of the ln HCE equation is a problem. The high value of the condition

[2] Simulation results by Leung and Yu (1996) showed that the two-part model predicts better than the Heckman model when multicollinearity is severe. This holds even if the Heckman model is the "true" model (see Jones, 2000, for an excellent survey on the debate).

[3] Seshamani and Gray (2004a) report a similar high R^2 for the same regression, using the ORLS data set.

number (see Greene, 1997) points in the same direction (see Table 3). This collinearity may explain why the age and other variables, including λ, are insignificant in the OLS estimation. We address this problem below by employing a two-part model that excludes the inverse Mill's ratio from the second equation.

The second criticism of the ZFM approach concerns potential endogeneity of regressors. In the present context, this means that causality does not only run from proximity of death to HCE as assumed but also in reverse, with higher HCE contributing to longer time to death. Salas and Raftery (2001) correctly pointed out that the OLS estimation is biased if the assumption of weak and strong exogeneity is violated. Here, weak exogeneity requires that proximity to death in a given month be not affected by HCE in a given month. A violation of this condition causes OLS coefficients to be biased and inconsistent, and their standard errors incorrect. Strong exogeneity requires that proximity of death is not affected by HCE of previous periods either. In the event that this condition is violated as well, Salas and Raftery (2001) proposed the following approach, claiming it to be equivalent to an instrumental variable (IV) estimation. Adding lagged HCE as omitted variables to the OLS equation serves to purge the error term of the components that cause the correlation with the explanatory variable, proximity to death in this instance (for details, see Kennedy, 1992, p. 148). Adding three such lagged variables still leaves a total of 19 months for representing proximity to death. One obtains (note that a higher value of j means increased distance from the time of death, i.e. an increasing lag)

$$\begin{aligned}\ln HCE_{im} | HCE_{im} > 0 = {} & \beta_0 + \beta_1 A_i + \beta_2 \frac{A_i^2}{1000} + \beta_3 S_i + \beta_4 (A_i \cdot S_i) \\ & + \beta_5 REG_i + \beta_6 \lambda_i + \sum_{j=1}^{19} \gamma_j M_{jm} \\ & + \sum_{n=m+1}^{m+4} \rho_k \ln HCE_{in} + \varepsilon_{im}. \end{aligned} \tag{3}$$

A test showing that the coefficients of the lagged variables are jointly equal to zero is equivalent to a "Hausman test" for equality of OLS and these quasi IV estimates. The left-hand side of Table 3 also contains the estimation results for regression. Clearly, the joint hypothesis of zero coefficients for all three lagged variables can be rejected. The assumption of strong exogeneity in the original OLS estimation is violated. However, all age-related variables continue to be insignificant. Again, the condition number is relatively large, pointing to multicollinearity in the regression. Moreover, this quasi IV approach is only valid if the lagged variables are uncorrelated with the error term. However, the error term may contain a time-invariant component reflecting permanent unmeasured characteristics of the individual. In this event, this orthogonality condition will be violated (see Kennedy, 1992). On the whole, the quasi IV approach suggests that the original ZFM finding that the proximity of death rather than age is a main determinant of HCE is fairly robust to endogeneity error.

4. Analysis of HCE in a given year as a function of time to death

Most of the published literature on the age effect on HCE of the deceased analyses the time path towards the time of death (ZFM, 1999; Seshamani and Gray, 2004a, 2004b;

Table 3:
Heckman two-step regression coefficient estimates for ln(HCE), 0 to 2 years prior to death: OLS and IV estimation (results corrected for heteroskedasticity)

Model	**ZFM specification; OLS**				**Augmented model; quasi IV**			
Dependent variable	**Probit**		**Ln(HCE)**		**Probit**		**ln(HCE)**	
Constant	−0.087	(0.134)	7.408**	(0.559)	−0.894**	(0.150)	2.145**	(0.398)
Age	−0.010**	(0.004)	−0.007	(0.006)	−0.004	(0.004)	0.009	(0.007)
(Age^2) /1000	0.211**	(0.025)	0.050	(0.066)	0.096**	(0.028)	0.003	(0.051)
Male	0.157	(0.093)	0.079	(0.139)	0.070	(0.103)	0.081	(0.194)
Age · Male	−0.005**	(0.001)	−0.003	(0.002)	−0.002	(0.001)	−0.003	(0.002)
Zurich	−0.260**	(0.020)	−0.029	(0.081)	−0.136**	(0.022)	−0.149**	(0.038)
Suppl. hospital insurance	0.027	(0.020)	−0.104**	(0.028)	0.029	(0.021)	−0.001	(0.039)
Lambda			−0.708	(0.577)			2.507**	(0.206)
M1[a]	0.576**	(0.058)	0.749**	(0.182)	0.276**	(0.063)	0.976**	(0.106)
M2	0.484**	(0.057)	0.505**	(0.161)	0.234**	(0.062)	0.697**	(0.106)
M3	0.383**	(0.056)	0.498**	(0.137)	0.158*	(0.062)	0.569**	(0.106)
M4	0.338**	(0.056)	0.433**	(0.127)	0.145*	(0.062)	0.528**	(0.106)
M5	0.281**	(0.056)	0.279**	(0.114)	0.094	(0.061)	0.319**	(0.107)
M6	0.264**	(0.056)	0.380**	(0.111)	0.125*	(0.061)	0.442**	(0.107)
M7	0.206**	(0.055)	0.243*	(0.099)	0.075	(0.061)	0.262*	(0.107)
M8	0.198**	(0.055)	0.201*	(0.097)	0.100	(0.060)	0.274*	(0.107)
M9	0.167**	(0.055)	0.181*	(0.092)	0.063	(0.060)	0.212*	(0.107)

M10	0.116*	(0.055)	0.176*	(0.084)	0.008	(0.060)	0.134	(0.108)
M11	0.075	(0.055)	0.141	(0.079)	−0.023	(0.060)	0.050	(0.108)
M12	0.131*	(0.055)	0.176*	(0.086)	0.052	(0.060)	0.222*	(0.108)
M13	0.096	(0.055)	0.098	(0.081)	0.022	(0.060)	0.127	(0.108)
M14	0.009	(0.055)	0.105	(0.076)	−0.090	(0.060)	−0.059	(0.110)
M15	0.084	(0.055)	0.153	(0.080)	0.021	(0.060)	0.182	(0.108)
M16	0.033	(0.055)	0.091	(0.077)	−0.038	(0.060)	0.013	(0.109)
M17	0.040	(0.055)	0.090	(0.077)	−0.018	(0.060)	0.067	(0.109)
M18	0.025	(0.055)	0.048	(0.076)	−0.015	(0.060)	0.009	(0.109)
M19	0.057	(0.055)	−0.002	(0.078)	0.036	(0.060)	0.074	(0.109)
ln(HCE-1)					0.078**	(0.003)	0.181**	(0.010)
ln(HCE-2)					0.079**	(0.003)	0.158**	(0.010)
ln(HCE-3)					0.104**	(0.003)	0.193**	(0.012)
Number of observations	21,900		13,802		21,900		13,802	
R^2 or Pseudo-R^2	0.126		0.0528		0.411		0.205	
Condition number			without λ	with λ			Without λ	with λ
			16	23.7			19.5	27

[a] *Note*: Month 20 (M20) is the reference left out dummy.
* Significant at the 95 per cent confidence level.
** Significant at the 99 per cent confidence level.

Stearns and Norton, 2004). These studies are confronted with potential endogeneity of the TTD dummy regressors, resulting in biased and inconsistent estimates. This distorts not only hypothesis testing but also forecasting. To circumvent this problem, we analyse HCE in a given period, the year 1999, as a function of the remaining time to death expressed in months. We estimate a two-part model (which treats the two equations as entirely independent) of the following form:

$$p(HCE_i > 0) = \beta_0 + \beta_1 A_i + \beta_2 \frac{A_i^2}{1000} + \beta_3 S_i + \beta_4 (S_i \cdot A_i) + \sum_{k=1}^{6} \beta_{5k} W_{ik} + \beta_6 D_i + \beta_7 (D \cdot A_i) + \beta_8 TTD_i + \varepsilon_i \quad (4)$$

$$HCE_i | HCE_i > 0 = \beta_0 + \beta_1 A_i + \beta_2 \frac{A_i^2}{1000} + \beta_3 S_i + \beta_4 (S_i \cdot A_i) + \sum_{k=1}^{6} \beta_{5k} W_{ik} + \beta_6 D_i + \beta_7 (D \cdot A_i) + \beta_8 TTD_i + \varphi_i \quad (5)$$

with $W = \{REG, ACC, HOSP, ODI, DED, EI\}$ where REG = regional dummy variable (Zurich = 1, Geneva = 0), *ACC*, *HOSP*, *OSI* and *DED* respectively are dummy variables for supplementary insurance (accident, hospital, other supplementary schemes) and optional high deductibles (with = 1 if the extras are chosen and = 0 otherwise), and *EI* is the average HCE level in the communities where the individuals live. This information comes from another source of data. Decedents are distinguished from survivors by the dummy variable D (= 1 if the individual died prior to the end of June 2003). Expected HCE of individual i then equals his or her probability of incurring costs times the conditional amount of HCE: $HCE_i = \Pr(HCE_i > 0) \cdot HCE_i | HCE_i > 0$.

Despite the log normal distribution of HCE data, we use arithmetic rather than logarithmic HCE here because this alternative allows a simple calculation of expected HCE and avoids the problems associated with the smearing factor if heteroskedasticity is present (Manning, 1998; Mullahy, 1998; Dow and Norton, 2002). We use the second set of data that includes both survivors and the deceased. Survivors are persons still alive by 30 June 2003. Their time to death is unknown by definition; however, it must exceed the maximum value of the deceased, which is 42 months. Therefore, $TTD = 43$ is coded for all survivors. In order to achieve a balanced ratio of deceased and survivors in every age class, we restrict the age range to (30, 95), since there are very few deaths among ages 30- and few survivors among ages 95+. In the constrained sample, there are 62,006 survivors and 3,657 deaths.

In principle, the equations to be estimated should not only contain the dummy variable for death but also interaction terms involving all age-related variables, on the one hand, and time to death (TTD), on the other. In this way, age and TTD could be pitted against each other as variables determining HCE both among the deceased and survivors. For example, if TTD itself were negative and significant while (TTD · Death) lacked significance, one would have to conclude that HCE decreases with TTD (increases with proximity to death) in the same way among the deceased and survivors. The original ZFM analysis, performed on the data of decedents only, could then be generalized to the entire population. However, including this interaction term in the regression caused perfect multicollinearity, quite likely because of the censored property of TTD. Including (Death · Age^2) had the same effect.

The estimation results pertaining to this incomplete specification are shown in Table 4. In the probit step, individuals who died during the observation period have a massively higher likelihood of positive HCE. All age-related variables are significant with expected signs, but so is TTD. In the equation for HCE with selection correction λ, all age-related variables are significant, in spite of a high condition number indicating collinearity. Again, TTD is a significant predictor, too. The inverse Mill's ratio is not significant in the second step of the Heckman estimation, which could be a consequence of collinearity or suggest that the two parts are indeed independent.

The two-part model is less plagued by multicollinearity, as expected. The same pattern of signs and levels of significance obtains. It may be noted that HCE of the deceased is estimated to be up to SFr 10,300 higher compared to survivors, which serves to underscore the importance of death (and, in the extreme, closeness to death) in the determination of HCE. Interestingly, individuals with supplementary hospital insurance appear to have a higher likelihood of positive HCE but not necessarily a higher level of HCE. Those having other supplements have a higher likelihood too, but their HCE is lower compared to the others. Finally, there is evidence that moral hazard effects are dampened by high deductibles, which are associated both with a lower likelihood of positive HCE and a lower conditional level of HCE.

Table 4:
Models of HCE with survivors and the deceased

Model						
Dependent variable	**Probit**		**Heckman model HCE**		**Two-part model HCE**	
Constant	0.173	(0.166)	14,429**	(1,242)	14,400**	(764)
Age	−0.003	(0.003)	−156**	(17)	−156**	(16)
(Age^2)/1000	0.163**	(0.027)	1,827**	(135)	1,826**	(135)
Male	−1.124**	(0.045)	1,570*	(703)	1,551**	(257)
Male · Age	0.013**	(0.001)	−29**	(9)	−29**	(4)
Death	0.909**	(0.182)	7,467**	(912)	7,482**	(780)
Death · Age	−0.011**	(0.002)	−62**	(11)	−62**	(10)
TTD	−0.008**	(0.003)	−234**	(10)	−234**	(10)
Zurich	−0.095**	(0.027)	−168	(150)	−168**	(147)
Suppl. hospital insurance	0.121**	(0.012)	−13	(84)	−12	(67)
Other suppl. schemes	0.274**	(0.023)	−1,415**	(174)	−1,412**	(130)
Accident insurance	−0.028*	(0.013)	680**	(78)	680**	(78)
High deductible	−0.328**	(0.012)	−750**	(159)	−754**	(67)
Average expenditure level	0.004**	(0.000)	21**	(3)	21**	(2)
Lambda			−32	(1,059)		
Number of observations	65,663		49,794		49,794	
R^2 or Pseudo-R^2	0.21		0.154		0.154	
Condition number			33		22.3	

** Significant at the 99 per cent confidence level.
* Significant at the 95 per cent confidence level.

As to the effect of all age-related variables combined, it is given by

$$\frac{\partial HCE(HCE > 0)}{\partial A} = \beta_1 + 2\beta_2 \frac{A}{1000} + \beta_4 S + \beta_7 \tag{6}$$

Applying a bootstrap to calculate standard errors based on the two-part model for conditional HCE,[4] we fail to find a significant combined effect among the deceased over a wide age range, i.e. $\partial HCE(HCE > 0)/\partial A = 0$. Only beyond age 80 is an additional year of age associated with an increase in HCE, in the magnitude of a modest SFr 100 (some U.S.$ 75).[5] The insignificance of the combined age effect may be puzzling in view of the significance of the individual coefficients in Table 4. However, these coefficients are of opposite sign, thus tend to compensate each other. Moreover, this compensating tendency increases with age because of Age^2. In combination, these effects may well result in a lack of significance. Hence, the findings of ZFM are again confirmed, by and large.

The predicted effect of age on HCE is illustrated in two figures. Figure 1 presents the estimated HCE for male survivors and the deceased, conditional on having positive HCE. The average TTD for the deceased is 21 months. HCE of the deceased is slightly U-shaped starting with 15,120 at the lower bound, with a minimum at age 65, and ending with 13,900 SFr at the upper bound of the age range. There, the curve has a small positive slope, reflecting the modest marginal effect indicated above. Note that at age 30 predicted conditional HCE of decedents exceeds that of survivors by a factor of almost four; it

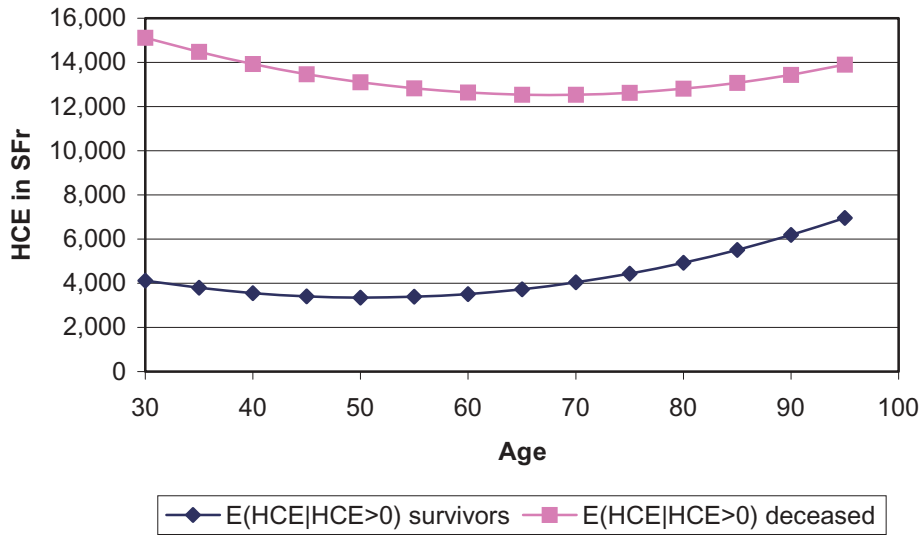

Figure 1: Conditional HCE of survived and deceased men

[4] We randomly draw 15,000 observations from the original sample and replicate the sample 100 times (procedure by LIMDEP 8.0).

[5] The effect is not significant, however. In fact, over the entire age range, the age gradient was insignificant for men. For women, the age gradient is only significant beyond age 80.

consistently falls to reach a value of 2.5 at age 95, pointing to a decreased relative intensity of medical treatment at higher ages.

The results regarding the very old contradict recent evidence suggesting that health resource utilization near the end of life is negatively age-dependent. Spillman and Lubitz (2000) with U.S. Medicare data, Felder *et al.* (2000) and Schellhorn *et al.* (2000) with Swiss data, and Chernichovsky and Markowitz (2004) with Israeli data, found a negative age gradient for elderly patients.

Among the survivors the combined effect of age on HCE is significantly positive for men beyond age 60 and women above age 50. For instance, men that survived in 1999 for at least 42 months incurred a conditional predicted HCE of SFr 3,540 if 30 years old but SFr 7,440 if 95 years old.

Figure 2 shows expected HCE of men as a function of age for survivors and the deceased, which calls for taking into account the likelihood of having positive HCE as well. For the deceased, the likelihood of positive HCE is around 0.9 (versus 0.65 for survivors), with a slight increase with age. This causes the U-shape of the curve to be a bit more marked towards the high end. Still, expected HCE of the deceased do not follow a monotonously rising trend, in keeping with the findings of ZFM. For the same reason, the upward trend for the survivors is reinforced as well, this time resulting in a consistently increasing trend. At age 30, predicted HCE of the deceased by far exceeds that of survivors again; here the multiple is almost six, falling to roughly two at age 95. Probably for the first time this constitutes evidence to the effect that younger decedents may be more costly both in absolute and relative terms than those in retirement age. Studies by Lubitz and Riley (1993) and Hogan *et al.* (2001), reporting factors between six and seven, are limited to Medicare data.

For comparison, Figure 2 also contains actual HCE of men (again in sample two). Not surprisingly, due to the high cost of the deceased, the curve of the actual HCE is substantially

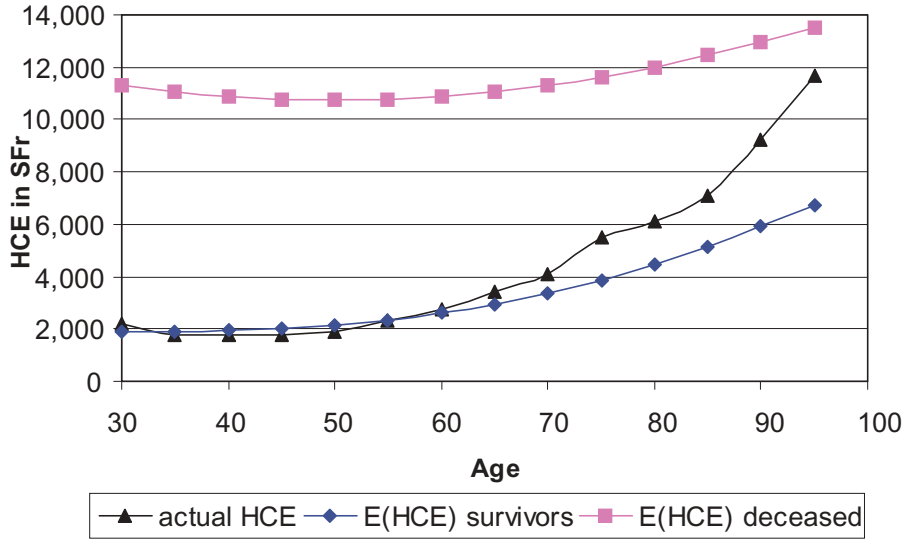

Figure 2: Expected and actual HCE of surviving and deceased men

steeper than the one of the survivors. Thus, a naïve estimation that does not consider the high cost of dying will overestimate by far the age effect on aggregate HCE. This upward bias can be evaluated as follows. First, note that the integral below the curve equals HCE incurred by the respective group (always beyond age 30). Next, integrating the curve E(HCE) for deceased men and women starting at age 30 yields the cost of dying; this amounts to 22 per cent of total HCE and can reasonably be claimed to be age-independent (also note that few deaths occur below age 30). This leaves 78 per cent of HCE beyond age 30 that potentially depends on age. Third, the same integration is performed on the E(HCE) curves of surviving men and women, netting out the benchmark expenditure, i.e. E(HCE) of 30-year-old men and women, respectively. Since these curves derive from a two-part model that controls for the proximity of death, their integrals provide an estimate of the share of total HCE that is truly attributable to increasing age, again beyond age 30. This figure amounts to 30 per cent. Finally, this integration can be applied to E(HCE) curves derived from a naïve two-part model that fails to control for proximity to death (by dropping TTD from all equations). Under this variant, 38 per cent of HCE is attributed to increasing age.

Bias in the estimate of dependence of HCE on age can now be expressed in several ways. If the standard of comparison is the conventional wisdom in its extreme form, 100 per cent of HCE beyond age 30 would be claimed to be potentially age-dependent. This figure shrinks to those 30 per cent that are truly attributable to age, resulting in an overestimate of more than a factor of three. An intermediate stance would be to acknowledge that among the deceased, HCE seems to be independent of age. The "true" age component of 30 per cent then has to be compared to the net potential component of 78 per cent, resulting in an overestimate by more than a factor of two. Finally, bias could refer to a model comparison. On this standard, the naïve model overestimates the effect of age by eight percentage points of HCE ($= 38 - 30$), more than one-quarter. Thus, whatever the standard of comparison used, these estimates point to substantial bias caused by neglecting the effect of the cost of dying on total HCE.

The estimates shown in Table 4 as well as the calculations based on them are subject to an important caveat. Recall that time to death essentially cannot be observed for survivors, necessitating the assignment of the fixed value $TTD = 43$ to the great majority of observations. This cannot but impart a substantial measurement error to TTD, causing its coefficient to be biased toward zero. Since the (negative) correlation between TTD and age-related variables is reduced as well, due to the construction of TTD, the standard errors of their coefficients are smaller than they would have otherwise been, creating a tendency towards statistical significance. In the comparison between time to death and age, this serves to tilt the balance in favour of age and against the red herring hypothesis.

5. Conclusion

Conventional wisdom suggests that a country's health care expenditure (HCE) rises with the proportion of elderly persons in the population. Predictions of future expenditure based on cross-section data tend to reaffirm this view. For instance Fuchs (1999a, 1999b) attributes the increase in the proportion of GDP spent on health in the U.S. to the "increase in the number of persons aged 65 and above due to the ageing of baby boomers".

At the aggregated level, however, there is no empirical evidence of a causal positive correlation between population ageing and HCE growth. Using OECD country data, Getzen (1992) and Barros (1998) found no evidence that population ageing drives HCE. For Getzen the conventional wisdom neglects the budget constraints confronting both governments and

private individuals that ultimately limit HCE. This dissent from conventional wisdom has been vindicated by recent research studying HCE of the deceased. When proximity to death is included as an explanatory variable, age tends to be statistically insignificant, suggesting that the positive relationship between age and average HCE reflects the high costs of dying and the high mortality in old age. In fact, Lubitz and Riley (1993) and Hogan *et al.* (2001) reported that Medicare payments per decedent are about six to seven times higher than those per survivor, a figure also found in HCE data of a Swiss sickness fund (Felder *et al.,* 2000). Probably for the first time this study points to a multiple of similar magnitude among individuals between age 30 and retirement age.

Summing up the evidence at the time, Zweifel *et al.* (1999) claimed population ageing to be a red herring in health care, distracting attention "from the choices that ought to be made", in particular with regard to the incentives to apply the newest medical technology even if the expected gains in life expectancy and quality of life are small. The case still stands. This paper reaffirms the original result of no age effect on HCE of the deceased over a wide range of ages. While it finds age-related variables to be significant in explaining HCE of the survivors, the age effect on aggregate HCE is much smaller when time to death is controlled for. Thus, a failure to take into account the high cost of dying definitely runs the danger of putting excessive emphasis on population ageing as a determinant of the future cost of health.

REFERENCES

BARROS, P.P., 1998, "The black box of health care expenditure growth determinants", *Health Economics,* 7, pp. 533-44.

DOW, W.H. and NORTON, E.C., 2002, "The red herring that eats cake: Heckit versus two-part model redux", *Triangle Health Economics Working Paper Series*, No. 1, University of North Carolina at Chapel Hill.

FELDER, S., MEIER, M. and SCHMITT, H., 2000, "Health care expenditure in the last months of life", *Journal of Health Economics,* 19, pp. 679-95.

CHERNICHOWSKE, D. and MARKOWITZ, S., 2004, "Ageing and aggregate cost of medical care: conceptual and policy issues" (forthcoming*, Health Economics*).

FUCHS, V.R., 1999a, "Provide: the economics of aging", A.J. Rettenmaier and T.R. Saving (eds), *Medicare Reform: Issues and Answers*. Chicago.

FUCHS, V.R, 1999b, "Health care for the elderly: How much? Who will pay for it?", *Health Affairs*, 18, pp. 11-21.

GETZEN, T.E., 1992, "Population ageing and the growth of health expenditure", *Journal of Gerontology: Social Sciences,* 47, pp. S98-S104.

GREENE, W.H., 1997, *Econometric Analysis*. London: Prentice-Hall.

HECKMAN, J.J., 1979, "Sample selection bias as a specification error", *Econometrica*, 47, pp. 153-62.

HOGAN, C., LUNNEY, J., GABEL, J. and LYNN, J., 2001, "Medicare beneficiaries' cost of care in the last year of life", *Health Affairs,* 20, pp. 188-95.

JONES, A., 2000, "Health Econometrics", in A.J. Culyer and J.P. Newhouse (eds), *Handbook of Health Economics*, Vol. 1. Amsterdam: Elsevier Science, pp. 265-343.

KENNEDY, K., 1992, *A Guide to Econometrics*. Oxford U.K. and Cambridge U.S.: Blackwell.

LEUNG, S.F. and YU, S., 1996, "On the choice between sample selection and two-part models", *Journal of Econometrics*, 72, pp. 197-229.

LUBITZ, J.B. and RILEY, G.F., 1993, "Trends in Medicare payments in the last year of life", *New England Journal of Medicine*, 328, pp. 1092-96.

MANNING, W.G., 1998, "The logged dependent variable, heteroskedasticity, and the retransformation problem", *Journal of Health Economics*, 17, pp. 283-95.

MULLAHY, J., 1998, "Much ado about two: reconsidering retransformation and the two-part model in health econometrics", *Journal of Health Economics*, 17, pp. 241-81.

O'NEILL, C., GROOM, L., AVERY, A.J., BOOT, D. and THORNHILL, K., 2000, "Age and proximity to death as predictors of GP care costs: Results from a study of nursing home patients", *Health Economics*, 9, pp. 733-38.

PUHANI, P.A., 2000, "The Heckman correction for sample selection and its critique", *Journal of Economic Surveys*, 14, pp. 53-68.
SALAS, C. and RAFTERY, J.P., 2001, "Econometric issues in testing the age neutrality of health care expenditure", *Health Economics Letters*, 10, pp. 669-71.
SCHELLHORN, M., STUCK, A.E., MINDER, C.E. and BECK, J.C., 2000, "Health services utilization of elderly Swiss: Evidence from panel data", *Health Economics*, 9, pp. 533-45.
SESHAMANI, M. and GRAY, A., 2004a, "Ageing and health care expenditure: the red herring argument revisited", *Health Economics*, 13, pp. 303-14.
SESHAMANI, M. and GRAY, A., 2004b, "A longitudinal study of the effects of age and time to death on hospital costs", *Journal of Health Economics*, 23, pp. 217-35.
SPILLMAN, B.C. and LUBITZ, J., 2000, "The effect of longevity on spending for acute and long-term care", *New England Journal of Medicine*, 342, pp. 1409-15.
STEARNS, S.C. and NORTON, E.C., 2004, "Time to include time to death? The future of health care expenditure predictions", *Health Economics*, 13, pp. 315-27.
VELLA, F., 1998, "Estimating models with sample selection bias: a survey", *Journal of Human Resources*, 33(1), pp. 127-69.
ZWEIFEL, P., FELDER, S. and MEIER, M., 1999, "Ageing of population and health care expenditure: a red herring?", *Health Economics*, 8, pp. 485-96.