© The Journal of Risk and Insurance, 1992 Vol LIX, No. 3, 443-460

Predictability of Individual Health Care Expenditures

René C. J. A. van Vliet*

ABSTRACT

It is widely believed in both academic and political circles that direct, uniform capitation payments to health insurers may induce them to contain costs. The premium-replacing capitation payments should account for predictable variations in individual health care expenditures. This article describes a model that encompasses and rejects several previous models that have been proposed for analyzing these variations. The results indicate that, at most, 20 percent of the variance among individuals in annual, short-term health care expenditures is predictable.

Introduction

As part of a move toward a more market oriented health care system in the Netherlands, the Dutch government is considering capitating competing health insurers on the basis of a method that resembles the Average Adjusted Per Capita Cost (AAPCC) method employed in the U.S. Medicare program for paying capitated delivery systems. An important consideration in developing this kind of capitation method is the choice of risk factors to be included in the capitation payment formula. If the risk factors chosen are not capable of adequately predicting future health care expenditures of individual insureds, two serious problems can be anticipated. First, funds may be distributed unfairly among the various insurance companies, with the result that insurers with a relatively "unhealthy" portfolio may go bankrupt. Second, insurers may select against those whose health care costs are predictably above their capitation payment.

In this context "adequate" prediction means that it should not be financially attractive for insurers to identify groups of people for which the capitation formula is either paying too much or too little. In other words, it

---

* René C. J. A. van Vliet is Associate Professor of Health Economics and Statistics in the Department of Health Care Policy and Management, Erasmus University, Rotterdam, the Netherlands. The author thanks Erik Schut, Wynand van de Ven, I and two anonymous reviewers for helpful comments on an earlier version of this article, and the health insurer Zilveren Kruis for providing the data set.

should be impossible for them to produce better cost predictions by gathering additional information.

Studies have been undertaken in the United States and in the Netherlands to identify additional risk factors that could be incorporated into the capitation formula (e.g., Ash et al., 1989; Homan, Glandon, and Counte, 1989; Rossiter, Adamache, and Faulknier, 1990; van Vliet and van de Ven, 1991). An essential question related to this research is whether ever-increasing refinements will ultimately lead to a capitation formula that predicts all cost differences between individuals. Intuitively, the answer must be no; health care insurance would be of little value if the future costs of individuals could be predicted with certainty beforehand.

The purpose of this study is to estimate an upper bound on the proportion of variance in annual individual health care expenditures that one may ever hope to predict. The results suggest that at least 80 percent of the variation in annual health care expenditures among individuals depends on chance and is thus fundamentally unpredictable. This finding implies that a capitation formula with a predictive accuracy far below 100 percent may be sufficient, making it unnecessary to gather additional data in an attempt to improve upon such a capitation formula. The results also provide a yardstick for comparing alternative capitation formulas: The closer the predictive ability is to 20 percent, the smaller are the above-mentioned problems of risk selection and unfair distribution.

Data

The analysis uses a panel data base comprising information over the five-year period 1976 through 1980 on the annual health care expenditures and insurance coverage of some 35,000 individuals who were continuously enrolled with Zilveren Kruis, the largest private health insurer in the Netherlands. The annual health care expenditures refer to short-term care and include the costs of inpatient room and board and of both inpatient and outpatient specialist care and ancillary services. The costs of drugs prescribed by physicians and care provided by a family doctor are excluded, because these are covered (and thus recorded by the insurance company) through supplementary insurance. During the period under study, insurers rarely offered deductibles or copayments. The data base contains information on age, sex, place of residence, and supplementary insurance coverage (for care provided by a family doctor or for treatment in a semiprivate hospital room). For a subgroup of about 14,000 people, additional survey data are available on indicators of health status in 1976.

The estimation of an upper bound on the proportion of variance in expenditures that could be predictable under the most favorable circumstances is based on plausible assumptions about the functional form of the autocorrelations of residual costs (i.e., after taking into account systematic effects of observed variables on total costs). Therefore, we first observe the correlations between individuals' actual expenditures in the five-year period (Table 1).

Table 1
Correlation Coefficients for Annual Health Care Expenditures by Year

	Year 1	Year 2	Year 3	Year 4	Year 5
Year 1	1	0.247	0.107	0.115	0.095
Year 2		1	0.211	0.127	0.119
Year 3			1	0.267	0.139
Year 4				1	0.247
Year 5					1

Note: The variances of the correlation coefficients (γ) can be approximated by 1/N. Since N = 35,000, this implies that 95 percent confidence intervals are given by γ +/- 0.010.

The autocorrelations of annual health care expenditures shown in Table 1 generally decrease with increasing time-lags, suggesting an autoregressive process with effects that taper off with time. The drop from first-order to second-order correlations is much sharper than those for longer time periods, which may be due to episodes of illness that result in bills that are reimbursed (but not necessarily incurred) in two successive years.

Models

McCall and Wai (1983), Beebe (1985, as cited by Welch, 1985), Welch (1985a), and Newhouse et al. (1989) found similar correlations in U.S. data. These studies propose three (error components) models that may explain the observed patterns—a variance components model (VC), an autoregressive model (AR), and a mixed autoregressive-variance components model (ARVC). With these models, it is possible to estimate the maximally predictable proportion of variance in health care expenditures. An alternative model is presented below, based on the autoregressive-moving averages models (ARMA), which are often employed in econometric analyses of time-series, cross-section data.

The idea that lies at the heart of previous models is that a person's health care expenditures each year are affected by the same personal and other characteristics, indirectly leading to correlations between expenditures in successive years. In this context one may distinguish between characteristics that will generally be known to both the insurer and the government agency responsible for the capitation payments—such as age, sex, insurance coverage, and place of residence—and characteristics that are difficult to observe, such as chronic health status. This distinction can be modeled with the following capitation formula:

$$Y_{i,t} = f(X_{i,t},\beta) + \varepsilon_{i,t}, \tag{1}$$

where $Y_{i,t}$ = health care expenditures of individual i in year t,
$X_{i,t}$ = a vector of observed characteristics that are assumed to affect expenditures via the function f and the parameters β, and
$\varepsilon_{i,t}$ = an individual and year specific error term with expectation 0.

Hence, $f(X,\beta)$ captures the systematic effects of the characteristics X, while the error term ε may be composed of two parts: strictly random and unpredictable fluctuations in costs (for instance, as the result of a minor car accident) or the effects of unobserved characteristics that are in principle predictable (chronic conditions, for example). The latter implies that the value of ε for an individual in a certain year may yield information on future costs. Various assumptions can be made with respect to the components of the error term ε. Combining these so-called error components models with the autocorrelations among residual costs estimated in a panel data base enables one to estimate the maximum proportion of variance in ε that is predictable (denoted by $maxR^2$). In this study, four of these models are considered. The remainder of this section describes in some detail an ARMA model and derives the relevant formulas. The formula derivations for the other three models, which are essentially special cases of the ARMA model, are analogous.

Because the term $f(X,\beta)$ is not of interest in the present context, it will be dropped in the subsequent exposition, and $y = Y - f(X,\beta)$ will be interpreted as the residual costs after correction for X (i.e., $y \equiv \varepsilon$). The index i is dropped to simplify notation. The vector X is assumed to contain the following risk factors: age, sex, insurance coverage, and place of residence. The ARMA model assumes that the residual y_t comprises three mutually uncorrelated components:

$$y_t = u_t + w_t + e_t, \tag{2}$$

$$u_t = u_{t-1}p + v_t \ (0 \leq p \leq 1), \tag{3}$$

$$w_t = re_{t-1} \ (0 \leq r \leq 1). \tag{4}$$

The term u_t can be interpreted as a permanent or structural component that, in part, persists over time and is caused primarily by the unobserved presence or absence of chronic conditions;[1] e_t is a transitory component (for example, due to acute, randomly occurring health problems); $w_t = re_{t-1}$ is caused by acute episodes of illness that commence at the end of year t-1 and for which part of the incurred costs are reimbursed and registered in year t; and v_t captures chronic conditions that arise in year t and extend to subsequent years. The components e_t and v_t are not correlated and have expectation 0 for each

---

[1] This term also comprises the impact of other structural factors that are not included in X, such as the regional supply of health services and the predisposition of people to use them. However, health status is very likely to be the most important omitted explanatory variable in equation (1).

individual; the variances of the various components—denoted by σ_y^2, σ_u^2, σ_w^2, σ_e^2, and σ_v^2—are assumed to be year independent. The model incorporates the most important aspects of the cost patterns observed in the previous section.

Two examples of future (residual) costs that may be known at the end of year t but are not incorporated in this model are the costs of pregnancy and medical treatments for which there are waiting lists. In addition, the ability to identify, on the basis of medical tests, individuals who are likely to develop serious diseases, is increasing (Office of Technology Assessment, 1988). The impact of this type of predictive information that cannot be inferred from past residual costs on overall predictability of health care costs is probably small.

Combining equations (2), (3), and (4) yields

$$y_t = u_{t-1}p + v_t + re_{t-1} + e_t. \tag{5}$$

So, the correlation between two successive years is

$$\mathrm{corr}(y_t\ , y_{t-1}) = p\sigma_u^2/\sigma_y^2 + r\sigma_e^2/\sigma_y^2 = p\rho + r\phi, \tag{6}$$

where $\rho = \sigma_u^2/\sigma_y^2$ and $\phi = \sigma_e^2/\sigma_y^2$. However, for $k > 1$,

$$\mathrm{corr}(y_t\ , y_{t-k}) = p^k\rho. \tag{7}$$

A comparison of equations (6) and (7) (the latter also being the relevant equation for the AR model) shows that for $k = 1$ an additional term is introduced.

The parameter ϕ can be derived from the other parameters by noting that, on the basis of equations (2) and (4),

$$\sigma_y^2 = \sigma_u^2 + (1 + r^2)\sigma_e^2, \tag{8}$$

and thus,

$$\phi = \sigma_e^2/\sigma_y^2 = (1 - \rho)/(1 + r^2). \tag{9}$$

At the end of year t the best conceivable predictor of y_{t+1} is

$$\hat{y}_{t+1} = pu_t + re_t. \tag{10}$$

If at that time u_t and e_t are available and p and r (both of which are independent of t) are known, then the proportion of variance in y_{t+1} that is predicted by $\hat{y}_{t+1}$ equals

$$\begin{aligned} \max R^2 &= 1 - [\mathrm{var}(y_{t+1} - \hat{y}_{t+1})/\mathrm{var}(y_{t+1})] \\ &= 1 - [(\sigma_v^2 + \sigma_e^2)/\sigma_y^2] = p^2\rho + r^2\cdot\phi, \end{aligned} \tag{11}$$

using the fact that v_t and e_t are uncorrelated.[2] Since there are no better predictors conceivable, $\max R^2$ in equation (11) can indeed be interpreted, under

[2] Note that the formula for R^2 applied in equation (11) is identical to the more conventional definitions—such as the squared correlation between y and $\hat{y}$, or the quotient of the variance of y and $\hat{y}$—only if $\hat{y}$ has been derived via an (explanatory) least squares regression. In that case, there holds that $(y-\hat{y})\perp\hat{y}$. Since the latter will not be true in general when making predictions, the more general definition of R^2 proposed by Efron (1978) is used (combined with the assumption that the mean of y is 0).

the assumptions of the model, as the maximally predictable proportion of variance in y_{t+1}. The ARMA model and the other three models utilized in this study are summarized in Table 2.

An alternative predictor for y_{t+1} can be obtained by employing y_t directly and performing a least squares regression:

$$\check{y}_{t+1} = (\bar{y}_{t+1} - \beta\bar{y}_t) + \beta y_t = \beta y_t, \tag{12}$$

where β is the first-order correlation (see equation [6]) and $\bar{y}_{t+1}$ and $\bar{y}_t$ are the averages of the respective residuals taken over all individuals, which are thus both equal to 0. The R^2 for this predictor is

$$R^2 = 1 - [\mathrm{var}(y_{t+1} - \check{y}_{t+1})/\mathrm{var}(y_{t+1})] = (p\rho + r\phi)^2. \tag{13}$$

Since p and r are in general (much) smaller than one, the R^2 value of equation (13) will be much smaller than that of equation (11). In other words, if this prediction procedure is employed, random fluctuations in y_t (in the form of the terms v_t and e_t) will result in a (much) smaller R^2 value than the maximum achievable. In the data base analyzed in this study, maxR^2 equals approximately 0.12, while the R^2 of equation (13) is only 0.05. These figures refer to predicting the residuals of the capitation formula of equation (1) and therefore must be combined with the R^2 of equation (1), which equals 0.021.

Table 2
Four Error Components Models

	VC	AR	ARVC	ARMA
Equations	$y_t=u_t+e_t$ $u_t=u_{t-1}$	$y_t=u_t+e_t$ $u_t=u_{t-1}p+v_t$	$y_t=u_t+w_t+e_t$ $u_t=u_{t-1}p+v_t$ $w_t=w_{t-1}$	$y_t=u_t+w_t+e_t$ $u_t=u_{t-1}p+v_t$ $w_t=re_{t-1}$
$\mathrm{corr}(y_t,y_{t-k})$	$\sigma_u^2/\sigma_y^2=\rho$	$p^k\rho$	$p^k\rho+\sigma_w^2/\sigma_y^2$	$p^k\rho+D_kr(1-\rho)/(1+r^2)$ ($D_1=1$, $D_k=0$ if $k\neq1$)
maxR^2	ρ	$p^2\rho$	$p^2\rho+\sigma_w^2/\sigma_y^2$	$p^2\rho+r^2(1-\rho)/(1+r^2)$

Note: VC = variance components model, AR = autoregressive model, ARVC = a mixed autoregressive-variance components model, and ARMA = autoregressive moving averages model. To simplify notation, the index i is dropped; y denotes the residual costs after correction for X (see also equation [1]).

The Models' Predictive Performance

The correlations among residual health care expenditures are shown in Table 3. The autocorrelation pattern in the residual costs is similar to that observed in Table 1 with respect to total costs. The autocorrelations in Table 3 are used to estimate the parameters of the four error components models (see Table 2). The parameter estimates are presented in Table 4.

Table 3
Correlation Coefficients for Residual Health Care Expenditures by Year

	Year 1	Year 2	Year 3	Year 4	Year 5
Year 1	1	0.233	0.091	0.101	0.076
Year 2		1	0.194	0.108	0.100
Year 3			1	0.253	0.121
Year 4				1	0.230
Year 5					1

Note: In the data base health care expenditures were regressed (using OLS) on age-sex (grouped into 18 x 2 classes), yes/no covered for the costs of family doctor care, yes/no covered for treatment in a semiprivate hospital room, and yes/no living in the western part of the country. This regression was performed on all person-years of data combined (N ≈ 5 x 35,000 = 175,000). The R^2 of this regression was 0.021. Next, the correlations among the resulting residuals were calculated.

Table 4
Estimated Parameters and Some Relevant Statistics

	VC	AR	ARVC	ARMA
$\rho=\sigma_u^2/\sigma_y^2$	0.151 (.022)	0.368 (.048)	0.999 (.90)	0.142 (.041)
p	—	0.601 (.054)	0.135 (.13)	0.873 (.096)
σ_w^2/σ_y^2	—	—	0.089 (.014)	—
r	—	—	—	0.123 (.025)
SSE[a]	0.043	0.0062	0.0026	0.0024
df.	9	8	7	7
test ARMA[b]	58.1 ($F_{2,7}$)	10.8 ($F_{1,7}$)	-3.85 & 1.81 (N)	—
maxR^2	0.151 (.022)	0.133 (.011)	0.108 (.011)	0.121 (.008)

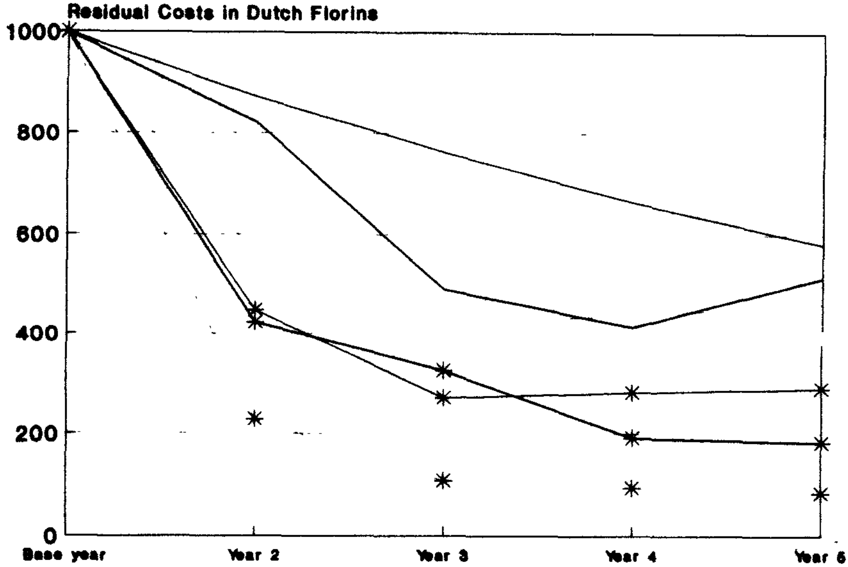
Note: VC = variance components model, AR = autoregressive model, ARVC = a mixed autoregressive-variance components model, and ARMA = autoregressive moving averages model. The parameters are estimated on the basis of the correlations presented in Table 3; standard errors are given in parentheses. The parameter of the VC model is estimated as the ratio of within-person variance and total variance using formulas provided by Searle (1971). The parameters of the other models are estimated by nonlinear least squares using the formulas for the correlations. In the estimation the variances of the correlations, which are very small (≈ 0.00003), are ignored. The so-called delta method is applied to obtain a standard error for the maxR^2 values as well as for the parameter r in the ARMA model (see, e.g., Bishop, Fienberg, and Holland, 1975, for a derivation of the delta method).

[a] SSE is the sum of squared differences between sample correlation coefficients and estimated coefficients based on the four models.

[b] By means of F-tests the null hypothesis is tested that the VC and AR models yield estimates that are equal to those produced by the ARMA model. The test for non-nested nonlinear models developed by Pesaran and Deaton (1978) was used to test the hypotheses that either the ARVC or the ARMA model is the correct model. This test yields statistics (elevated by N) that have asymptotic standard normal distributions.

The performance of the models is illustrated in Figure 1, which displays the sample correlations of Table 3 along with the estimated values based on the four models. The ARMA model is more capable of representing the sharp drop from first- to second-order correlations than are the VC and AR models. The VC model predicts constant correlations based on the unrealistic assumption that the structural component u_t of the error term does not change over time. The main differences between the models emerge in the long run: The AR model predicts ever-decreasing correlations, the ARVC model predicts constant correlations after about four or five years, and the ARMA model takes an intermediate position.

Figure 1
Sample Correlations and Estimated Correlations
Based on Four Error Components Models



Note: This figure displays correlations between residual costs of individuals in a certain base year (year 0) and residual costs in the subsequent ten years. The sample correlations are from Table 3. The estimated correlations are calculated using the parameters of Table 2 and the equations of Table 4.

The VC model correlates poorly with the sample, as indicated by the sum of squared differences (SSE) presented in Table 4. The null hypothesis that the VC model yields estimates that are as accurate as those of the ARMA model is rejected decisively (F = 58.1). The corresponding null hypothesis for the AR model indicates that the ARMA model also reflects the observed correlations better than the AR model (F = 10.8). The differences between the ARMA and the ARVC models are small, but the null hypothesis that the ARVC model is a better than the ARMA model is rejected (N = -3.85).[3] The reverse hypothesis, that the ARMA model is more accurate than the

[3] The null and alternative hypothesis of this test are interchangeable. Therefore, it is possible that both hypotheses are rejected, implying that both models poorly represent the observed phenomena.

ARVC model, cannot be rejected ($N = 1.81$).[4] The main reason for the rejection of the VC and AR models is the large sample size combined with a relatively long time period. In other studies, data limitations prevented the analysis of more complex models (Newhouse et al., 1989; Welch, 1985a).

In sum, the VC and AR models both have poor fits compared with the other models. The differences between ARVC and ARMA are small, but based on theoretical grounds, the test results, and the greater standard error of $\max R^2$ for the ARVC model, the ARMA model is preferable. Therefore, the remainder of this article uses the latter model.

Predictive Variance

The $\max R^2$ values for the four models range from 0.108 to 0.151. The $\max R^2$ for the ARMA model is 0.121, implying that 12.1 percent of residual costs are predictable. Combining this with the R^2 of 0.021 for the capitation formula of equation (1), it is evident that, for this sample, at most, 13.9 percent ($0.021 + (1 - 0.021) \times 0.121$) of the variance in health care expenditures among individuals is in principle predictable.

The predictable variance of 12.1 percent in the residuals is composed of 10.8 percent ($p^2\rho 100$) as a result of the structural component in residual costs, while 1.3 percent ($r^2\phi 100$) is caused by transitory health problems for which costs are reimbursed in two consecutive years. The 87.9 percent of the residual variance that cannot be predicted is composed of 84.5 percent ($\phi 100$) due to strictly random fluctuations, while the remaining 3.4 percent is attributable to the manifestation of new illnesses that will generate additional costs in years to come.

The ARMA model can be used to resolve the issue of the predictability of the term e_t in the AR model. Newhouse (1986) argues that the AR model is inadequate in this respect because in many instances transitory illnesses can be predicted, for example, in the case of a pregnancy or an acute condition that will require treatment in the next period. Consequently, e_t and e_{t-1} in the AR model may be correlated (cf. Table 2). If the ARMA model is correct, this correlation is given by $r(1-\rho)/(1+r^2)$. Using the estimated parameters given in Table 4, this expression equals 0.104 with a standard error of 0.025, lending support to the hypothesis that at least part of the transitory illness component (e-term) in the AR model is predictable.

McCall and Wai (1983), Anderson and Knickman (1984), Beebe (1985, as cited by Welch, 1985), and Newhouse et al. (1989) found comparable sample correlations for completely different populations and for different time periods. The predictable variance that can be deduced from these sample correlations are around 14 percent and thus correspond closely with the 13.9 percent obtained above. Studies that attempt to explain annual

[4] These non-nested models were also combined in a variant that includes both models as special cases. This variant, however, produced a SSE equal to that of the ARMA model, indicating that the added complexity did not add accuracy.

health care costs of individuals based on individual and other characteristics have rarely been able to explain more than 20 percent of health care cost variability (Newhouse, 1986; van Vliet, 1988).[5] It seems reasonable to conclude that this article's findings are valid beyond the sample and study period used here.

Table 5
Correlation Coefficients for Residual Expenditures by Type of Expenditure[a]

	Year 1	Year 2	Year 3	Year 4	Year 5
Inpatient care					
Year 1	1	0.207	0.071	0.082	0.059
Year 2		1	0.158	0.077	0.071
Year 3			1	0.230	0.097
Year 4				1	0.193
Year 5					1
Outpatient care					
Year 1	1	0.304	0.218	0.182	0.136
Year 2		1	0.293	0.228	0.177
Year 3			1	0.330	0.228
Year 4				1	0.289
Year 5					1

[a]The R^2-values for eq. (1) for inpatient and outpatient care amount to 0.016 and 0.025 respectively The max R^2-values that can be calculated from this table using the ARMA model are 0.095 and 0.226, so that maximally 10.9 percent and 24.5 percent of the variance in these cost types is predictable.

Inpatient vs. Outpatient Costs

Some types of health care expenditures are easier to predict than others. Table 5 compares the autocorrelations of residual costs for inpatient and outpatient care. The correlations for inpatient care are substantially smaller, probably because of greater fluctuations in individual inpatient care from year to year. Maximally, 10.9 percent of the variance in inpatient costs and 24.5 of the variance in outpatient costs appear to be predictable at the individual level. The finding that inpatient costs are less predictable than outpatient costs corresponds with the results of previous studies (for instance, Newhouse et al., 1989).

The latter conclusion suggests that the maximally predictable proportion of variance in individual health care expenditures may change if more or other types of expenditures are included in the analysis. So far, the cost figures have not included the costs of care provided by a family doctor and the costs of prescription drugs, because not everyone in the sample had this

[5] Based on R^2 values adjusted for degrees of freedom.

type of supplementary coverage. For the subgroup of people who were covered for these costs--and whose costs are consequently known to the insurer--the maxR^2 for residual cost including these types of expenditures amounted to 0.150. The R^2 for capitation formula (1) equaled 0.0266, so that maximally 17.3 percent (0.0266 + (1 – 0.0266) x 0.150) of the variance in total costs is predictable, as compared with 13.9 percent if these types of expenditures are excluded. This higher percentage is mainly due to the high predictability of the expenditure types in question: 80 percent of the variance in the costs of care provided by a family doctor and the costs of prescribed drugs is predictable in this group.

Predictability Increases with Age

Finally, one may expect that the predictability of health care costs increases with an individual's age, as chronic illness becomes more likely. In order to test this hypothesis, the predictability was estimated separately for those in the sample over age 50 and for those under age 30. The maximally predictable proportion of variance in individual health care expenditures for these groups are 0.125 and 0.109, respectively. Combined with the R^2 values of 0.011 and 0.008 for the capitation formula in equation (1), we find that up to 13.4 and 11.6 percent of the cost variances in these groups are predictable. Although the difference in predictability seems small, the underlying parameters of the ARMA model for the two groups show that the drop in autocorrelations when time lags increase is much slower for the older population. This supports the hypothesis that predictability increases with age and suggests that differences in health care costs for older individuals are more predictable than those for young people.

In conclusion, in the analyzed data base at most 14 percent of the differences among individuals in annual health care costs (inpatient care and outpatient specialist care) can be predicted. When the costs of family doctor care and the costs of prescription drugs are included in the analysis, predictability rises to about 17 percent. Combining these findings with the results of other studies, one may reasonably conclude that, at most, 20 percent of the variance in annual short-term health care expenditures is predictable at the individual level.[6] By far, the majority of the differences in costs among individuals are random and therefore unpredictable.

---

[6] The assumption is that the systematic patterns in health care expenditures remain largely unchanged over time. In principle, the predictable part of expenditures may rise when the efficiency of health care provision improves, when more outpatient care is substituted for inpatient care, when more people take a (high) deductible, or when the costs of long-term nursing home care and care for the elderly are also financed through the capitation formula. Particularly, the latter would result in a substantial increase in maximally achievable predictability, while a crude capitation formula that also incorporates these types of long-term care but that does not take institutional status into account will have an even poorer R^2 than that for the short term care expenditures analyzed in this article, thereby increasing the potential profits of risk selection.

Applications

The capitation formula of equation (1), which uses as risk factors age, sex, insurance coverage, and place of residence, predicts only 2.1 percent of the variability in individuals' annual health care expenditures. This is a small fraction of the maximum predictable proportion of 13.9 percent. With information only on residual costs in a given year, an estimated 7 percent of the cost variation in the following year could be predicted (see equation [13]). What are the financial consequences of these and other differences in predictability? This section illustrates some of these consequences using six examples.

Example 1: Regression Toward the Mean

Assume first that the (residual) costs in year t are known. Analogous to equation (12), it may be verified that the estimated correlations of Figure 1 can be interpreted as the coefficients of least squares regressions of residual costs in year t on residual costs in year t-k ($k \geq 1$). Using the estimated correlations, the ARMA model predicts that people whose costs are in a certain year 1,000 Dutch florins above their expected costs as calculated with equation (1) will have costs in the next year that exceed expected costs by Dfl 228 (1,000 x 0.228; cf. Figure 1).[7] For the three subsequent years, this amount is Dfl 108, Dfl 94, and Dfl 84, respectively. This illustrates the well-known phenomenon of "regression toward the mean"; that is, if people are grouped according to their (residual) costs in a certain base year, the group means have a tendency to move toward the total mean over time. These numbers may seem small in comparison with the initial deviation of Dfl 1,000. However, aggregated over the four-year period, the extra costs amount to Dfl 512. These findings illustrate the effects of the difference in predictive accuracy between the capitation formula (which in fact predicts zero extra costs) and a model that also takes into account last year's costs.

Example 2: A Benchmark Prediction

Suppose now that for every individual the structural component u_t (caused by the unobserved presence or absence of a chronic condition) of the residual costs in year t (y_t) is known. The estimated extra costs in the following four years for a person for whom this component equals Dfl 1,000 amount to Dfl 873, Dfl 762, Dfl 655, and Dfl 581, respectively.[8] Aggregated, this amounts to Dfl 2,881. This may be seen as a benchmark: Better predictions seem impossible, so the predictive abilities of other risk factors

[7] The cost figures in this section are in Dutch florins of 1989, when one U.S. dollar was worth about Dfl 2, and mean short-term health care expenditures for privately insured individuals were about Dfl 1,078.

[8] The regression coefficient for the least squares regression of y_{t+k} on u_t equals $\mathrm{cov}(y_{t+k}, u_t)/\mathrm{var}(u_t)$, which in the ARMA model amounts to p^k (for $k \geq 1$; cf. equation [7]).

can be compared with this benchmark. For example, we may compare the latter figure of Dfl 2,881 with the extra costs of Dfl 512 that were predicted if only residual costs were available, as in example 1.

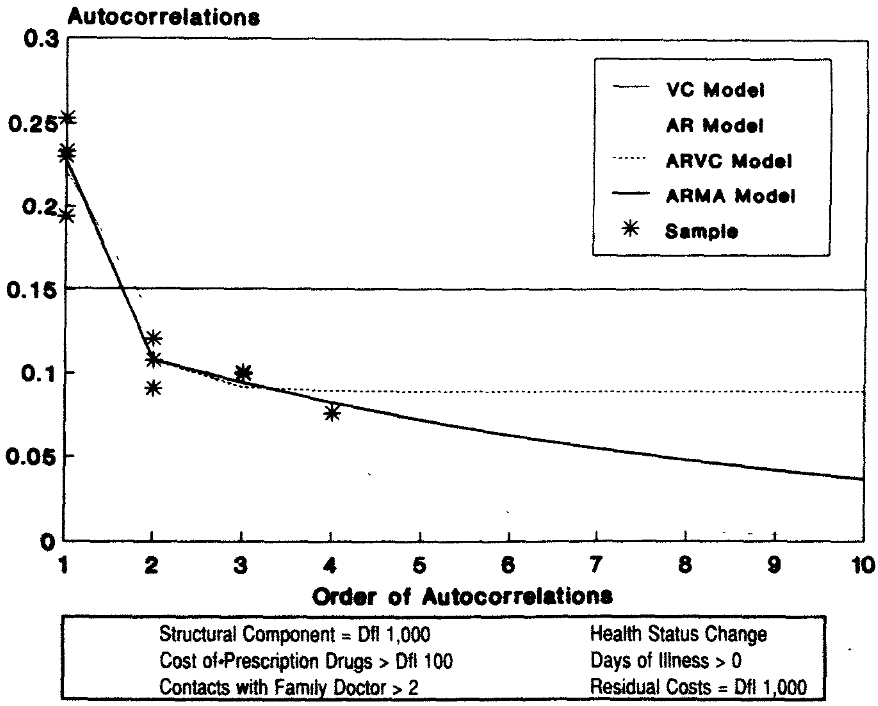
Example 3: Proxies for u_t

In practice one cannot know u_t, but good proxies may be available. For part of the analyzed data base (N ≈ 14,000), information is available on some health status indicators and medical consumption variables from a survey taken in 1976. Table 6 presents mean residual costs for several subgroups formed on these variables. Figure 2 displays the results of this analysis together with the findings of examples 1 and 2.

Figure 2 shows that people who were ill in the first half of the base year had health care expenditures in the following four years that were respectively Dfl 423, Dfl 325, Dfl 192, and Dfl 183 above their expected costs.

Figure 2

Future Residual Costs for People Grouped on Information in the Base Year (1976)



Note: Residual costs in years two through five for people grouped on their structural component or residual costs in the base year are estimated in examples 1 and 2. The mean residual costs in years two through five for the other groups are defined as the difference between actual and expected costs, the latter calculated on the basis of age, sex, insurance coverage, and place of residence. See also the footnotes of Table 6.

These numbers are substantially higher than the corresponding numbers for the group formed on residual costs in the base year (example 1). For the group with two or more contacts with their family doctor during the first half of 1976, the numbers are higher still. However, expenditures on prescribed drugs is an even better predictor, with extra costs in years 1977 through 1980 of Dfl 822, Dfl 490, Dfl 414, and Dfl 513, respectively, for the group with expenses above Dfl 100 in 1976. These numbers are only slightly lower than those for someone whose structural component u_t equals Dfl 1,000 in the base year. In other words, expenses for prescribed drugs give a good prediction of future, total health care expenditures. This is probably because of the strong correlation between drug use and the presence of chronic illnesses.

These examples are by no means extraordinary, as the considerable size of the various subgroups indicates (see Table 6). These findings suggest

Table 6
Future Residual Costs for Various Subgroups in Case of a Crude Capitation Formula[a]

	Year 1		Residual = Actual Costs - Expected Costs			
Risk Factor and Interval[b]	**N (%)**	**Mean Residual in Dutch Florins**	**Year 2**	**Year 3**	**Year 4**	**Year 5**
Health Status Relative to Last Two Years[c]						
Unchanged	85.9%	–310	–128	–70	–85	–94
Worse/Better	14.1	1941	855	518	538	555
Days of Illness (First Half of Year 1)						
0	73.4%	–516	–206	–165	–126	–129
> = 1	26.6	1391	588	452	267	255
Family Doctor Consultations (First Half of Year 1)						
0	46.9%	–338	–235	–157	–144	–178
1-2	33.9	–151	–25	–23	–98	–75
≥ 3	19.2	1041	581	440	456	478
Costs of Prescription Drugs in Dfl (First Half of Year 1)[d]						
0	56.7%	–389	–285	–228	–198	–212
1-100	21.3	–102	–100	4	–28	–128
≥ 101	21.9	976	802	478	403	501

[a] Residual costs are in Dutch florins of 1989, when one U.S. dollar was worth about 2 Dutch florins, and mean health care expenditures for privately insured individuals were about Dfl 1,078.
[b] To facilitate comparisons with the examples in the text, the class boundaries were chosen so that the mean residual costs in the base year (1976) for the high cost groups were approximately Dfl 1,000. For Figure 2, the resulting numbers have been rescaled to make these means exactly equal to Dfl 1,000.
[c] Health status in the first half of the base year relative to that in the preceding two years (self-rated).
[d] The costs of prescription drugs are also in Dutch florins of 1989.

that it is possible to obtain good estimates of the structural component of (residual) costs by means of various proxies of (chronic) illness.

Until this point it was assumed that data on costs and risk factors are available for a certain base year and that from these, costs in the following years had to be estimated. In general, however, data may be available for earlier years and longer future predictions may be needed. (The proposals of the Dutch government for capitating health insurers indicate that the mandatory contract period may become two years.)

Example 4: Past Averages Do Not Predict Future Spending

Assume therefore that we want to predict costs one year in advance (y_{t+1}) using average residual costs in previous years. The required expression for one year is given in equation (13). If data are available for years 1 to t, then the predictor is given by

$$\hat{y}_{t+1} = \Sigma\, y_j/t, \tag{14}$$

and the corresponding R^2 is given by

$$R^2 = [(\Sigma\, p^j)\rho + r\phi]^2 \cdot \sigma_y^2/[\mathrm{var}(\Sigma\, y_j)], \tag{15}$$

with

$$[\mathrm{var}(\Sigma\, y_j)] = \sigma_y^2\{t + 2(t\text{-}1)r\phi + 2\rho[\Sigma\, (t\text{-}j)p^j]\}. \tag{16}$$

Substituting the estimated parameters from Table 4 into equation (15) reveals that the predictable variance of health care expenditures decreases from 0.052 for t = 1 to 0.034 for t = 10. Thus, it is not sensible to employ average past expenditures to predict future costs. This is due to the autoregressive nature of health care expenditures, whereby expenditures of any two years are less correlated the longer the time lag between these years.

Example 5: Far Past Data Are Useless

It seems more sensible to apply a weighing method so that costs from the more distant past are given a lower weight. This clearly resembles the problems addressed by credibility theory (Goovaerts and Hoogstad, 1987). A simple weighing method can be obtained by regressing y_{t+1} on y_1 to y_t. The least squares coefficient vector b of this regression is given by

$$b = C^{-1}c, \tag{17}$$

where C denotes the correlation matrix of the residuals y_1 to y_t, and c is the vector with correlations between y_{t+1} and y_1 to y_t. The R^2 for this regression amounts to

$$R^2 = c'b. \tag{18}$$

Combining equation (17) with the estimated parameters of Table 3 shows that the weight of y_t in the vector b is at least three times as great as those of the residuals from more distant years, while the weights steadily decrease with the passing of time. The R^2 of equation (18) increases slowly from 0.052 for t = 1 to 0.062 for t = 5 and to 0.064 for t = 10. Thus, cost data from longer periods than the past one or two years are practically useless for making cost predictions at the individual level.

Example 6: Ten-Year Contract Period

Assume finally that information is available on u_t and e_t and that predictions are required for years t + 1, t + 2, etc. The maximum explainable variance in y_{t+1} amounts to 0.121. For year t + k (k > 1) the corresponding $\max R^2(k)$ is, by analogy with equation (11), given by

$$\max R^2(k) = p^{2k}\rho. \tag{19}$$

The value of $\max R^2(k)$ decreases from 0.082 for k = 2 to 0.036 for k = 5 and 0.009 for k = 10. This leads to the conclusion that after a period of ten years knowledge about u and e in the base year has virtually no predictive value. In other words, if the crude capitation formula of equation (1) was employed, then the length of the contract period should be *at least* ten years to prevent the problems of unfair distribution and risk selection. During this ten-year period these problems could, of course, still be considerable, as shown in example 3. Contract periods of this length do not seem very realistic, because they would seriously frustrate consumer choice and thereby reduce competition among insurers.

Conclusions

The ARMA model described in this article provides a plausible and empirically supported explanation of observed patterns in autocorrelations among individuals' annual health care expenditures. This model estimates the maximally predictable proportion of the variance in annual expenditures at the individual level. In the analyzed sample, approximately 14 percent of the differences in (inpatient and specialist outpatient) costs among individuals were predictable. If the expenditures for care provided by a family doctor and prescription drugs are taken into account, about 17 percent of the cost differences are predictable. In conjunction with the findings of other studies, one might reasonably conclude that no more than 20 percent of cost variations are predictable at the individual level.

A crude capitation formula based on age, sex, insurance coverage, and place of residence yielded an R^2 value of 2.1 percent, only about one-seventh of the maximally predictable amount. The financial consequences of this difference are considerable. An individual whose costs in the base year (1976) were Dfl 1,000 higher than expected due to structural factors (see example 2) would accrue costs in the ensuing four-year period that were Dfl 2,881 higher than expected based on the crude capitation formula. Incorporating the prior year's costs as an additional risk factor raises the R^2 value of this capitation formula from one-seventh to one-half of its maximally achievable value. Moreover, if cost data from the past one or two years are incorporated in the formula, the addition of cost data from more distant years yields little improvement in predictive accuracy. Various types of past expenditures, such as prescription drug purchases, appear to provide additional explanatory power. The findings also suggest that the predictive power of the structural component of (residual) costs in the ARMA model may be approximated by various proxies for chronic health status.

Discussion

In many countries (for example, Australia, Germany, Israel, New Zealand, Switzerland, and the United Kingdom) the concept of capitating health care insurers is receiving increasing attention, because it is believed that capitation may induce insurers in a competitive environment to concentrate more on cost-containment. The necessary premium-replacing capitation payments are supposed to be risk dependent; that is, they should account for systematic, predictable variations in health care expenditures. Therefore, it is important to investigate both the predictive accuracy of various alternative capitation formulas and the extent to which variations in expenditures in fact could be predictable.

This study shows the poor predictive performance in the Dutch health care system of a capitation formula based on age, sex, insurance coverage, and place of residence. It indicates that the difference between the predictive ability of this crude capitation formula and the maximum achievable predictive ability could have serious consequences. These unintentional consequences could be of a financial nature if insurers with relatively unhealthy portfolios are underpaid and those with healthy portfolios are overpaid. Also, risk selection by insurers may occur if potential profits outweigh selection costs.[9]

These problems may be mitigated by incorporating into the capitation formula additional risk factors, such as the costs incurred by an individual in the past. This has been shown in many studies to be a relatively good predictor of future costs. Cost information from only the past one or two years is necessary, because adding cost data from earlier years does little to improve predictability. However, the predictive ability of prior costs amounts to only about half of the maximum predictability achievable, possibly leading to the financial and risk selection problems mentioned above. Moreover, incorporating prior costs in the capitation formula may lead to strategic behavior by insurers (inflating costs in order to get higher capitation payments in the future) and new forms of risk selection (an

---

[9] Van de Ven and Van Vliet (1991) give an extensive overview of subtle strategies that insurers might use to select risks, despite the open enrollment proposed in the Dutch health care reform plans. In a competitive insurance market, risk selection may be necessary in order to achieve a stable market equilibrium (Schut, 1991). However, in the envisaged *regulated* Dutch health insurance market with open enrollment periods, it may become a destabilizing factor.

insured with a high-cost, self-limiting illness last year yields a high capitation payment but low costs this year). Risk factors that are more promising predictors include prior utilization combined with diagnostic information, disability, functional status, and indicators of chronic medical conditions (cf. van de Ven and van Vliet, 1991). Apart from having good predictive ability, these risk factors score well on other important criteria, such as validity, reliability, invulnerability to manipulation, obtainability, and absence of perverse incentives. However, much research is needed before these risk factors could be applied in the Dutch situation.

Finally, it should be noted that a capitation formula with an R^2 that is somewhat lower than the maximum achievable may very well suffice in practice, provided additional precompetitive measures are taken (cf. van de Ven and van Vliet, 1991). This seems to contradict the rather pessimistic findings by Newhouse et al. (1989), who showed that even a small difference between potential and actual explained variance may result in substantial profits if the insurers are able to find the "right" groups of insureds. The question is, however, what methods insurers have to locate these groups. Especially when all the easily available risk factors are included in the capitation formula, the selection costs may well outweigh potential profits. Together with precompetitive measures, this may induce insurers to abstain from risk selection.

In addition, some risk factors that are known to affect health care expenditures, such as supply and prices of health services, probably should not be accounted for at all in the capitation payments, because they are supposedly the health insurers' responsibility. This may be an additional reason for being satisfied with a lower than maximally achievable predictive accuracy of the capitation formula.

References

Anderson, G. F. and J. Knickman, 1984, Patterns of Expenditure Among High Utilizers of Medical Care Services: The Experience of Medicare Beneficiaries from 1974 to 1977, *Medical Care*, 22: 113-119.

Ash, A., F. Porell, L. Gruenberg, E. Sawitz, and A. Beiser, 1989, Adjusting Medicare Capitation Payments Using Prior Hospitalization Data, *Health Care Financing Review*, 10(4): 17-29.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland, 1975, *Discrete Multivariate Analysis: Theory and Practice* (Cambridge, Mass.: MIT Press).

Efron, B., 1978, Regression and ANOVA with Zero-One Data: Measures of Residual Variation, *Journal of the American Statistical Association*, 73(1): 113-121.

Goovaerts, M. J. and W. J. Hoogstad, 1987, Credibility Theory, Survey of Actuarial Studies, Number 4, Nationale Nederlanden, Rotterdam.

Homan, R. K., G. L. Glandon, and M. A. Counte, 1989, Perceived Risk: The Link to Plan Selection and Future Utilization, *Journal of Risk and Insurance,* 56: 67-82.

McCall, N. and H. S. Wai, 1983, An Analysis of the Use of Medicare Services by the Continuously Enrolled Aged, *Medical Care,* 21(6): 567-585.

Newhouse, J. P., 1986, Rate Adjusters for Medicare Under Capitation, *Health Care Financing Review,* Annual Supplement: 45-55.

Newhouse, J. P., W. G. Manning, E. M. Keeler, and E. M. Sloss, 1989, Adjusting Capitation Rates Using Objective Health Measures and Prior Utilization, *Health Care Financing Review,* 10(3): 41-54.

Office of Technology Assessment, 1989, *Medical Testing and Health Insurance,* OTA-H-384 (Washington, D.C.: Congress of the United States).

Pesaran, M. H. and A. S. Deaton, 1978, Testing Non-nested Nonlinear Regression Models, *Econometrica,* 46: 677-694.

Rossiter, L. F., K. W. Adamache, and T. Faulknier, 1990, A Blended Sector Rate Adjustment for the Medicare AAPCC When Risk-Based Market Penetration Is High, *Journal of Risk and Insurance,* 57: 220-239.

Schut, F. T., 1991, Equilibria in Competitive Insurance Markets When Risk Classification Is Not for Free, Unpublished report, Department of Health Care Policy and Management, Erasmus University Rotterdam, the Netherlands.

Searle, S. R., 1971, *Linear Models* (New York: John Wiley).

van de Ven, W. P. M. M. and R. C. J. A. van Vliet, 1991, How Can We Prevent Cream Skimming in a Competitive Health Insurance Market? The Great Challenge for the 90s, in: P. Zweifel and T. Frech, eds., *Health Economics Worldwide* (Amsterdam: Kluwer).

van Vliet, R. C. J. A., 1988, Hospital Utilization, Performance Measures and Health Status, Unpublished doctoral dissertation, Erasmus University Rotterdam, the Netherlands.

van Vliet, R. C. J. A. and W. P. M. M. van de Ven, 1991, Towards a Capitation Formula for Competing Health Insurers: An Empirical Analysis, *Social Science and Medicine,* 34: 1035-1048.

Welch, W. P., 1985, Regression Towards the Mean in Medical Care Costs: Implications for Biased Selection in HMOs, *Medical Care,* 23: 1234-1241.

Welch, W. P., 1985a, Medicare Capitation Payments to HMOs in Light of Regression Toward the Mean in Health Care Cost, in: R. M. Scheffler and L. F. Rossiter, eds., *Advances in Health Economics and Health Services Research* (London: JAI Press).