HEALTH-BASED RISK ADJUSTMENT IN SWITZERLAND: AN EXPLORATION USING MEDICAL INFORMATION FROM PRIOR HOSPITALISATION

Alberto Holly, Lucien Gardiol, Yves Eggli,

Tarik Yalcin and Tiago Ribeiro

First Version of the Final Report: November 2003

Revised Final Version: August 2004

Research financed by the Swiss National Fund,
National Research Program 45 "Future Problems of the Welfare State"
Grant No. 4045-059720/1

Institut d'économie et management de la santé (IEMS)
University of Lausanne
Champ de l'Air
Bugnon 21
1005 Lausanne
Switzerland

FOREWORD

This revised version of the first version of the final report greatly benefited from the remarks and suggestions of Prof. Tiago Ribeiro (IEMS). In particular, his help in the revision of Chapters 7 and 8 amounted to co-authorship. This is why his name appears in the list of the authors of this modified version of the final report.

ACKNOWLEDGEMENT

The authors are grateful to the Swiss National Fund (SNF) for its financial support.

The authors would like to thank the following sickness insurance funds:

CSS Assurance, HELSANA, and KPT/CPT

as well as the Centre Hospitalier Universitaire Vaudois (CHUV), and the University Hospital of Zurich (UniversitätsSpital – Zurich)

for providing the data used in this study.

In particular, they would like to thank the following persons for their help:

Dr. Konstantin Beck, CSS Assurance

Ute Kunze, CSS Assurance

Timo Widmer, HELSANA

François-Xavier de Rossi, KPT/CPT

Christiane Roth (MD), Director UniversitätsSpital – Zurich

Alfred Bollinger (MD), UniversitätsSpital – Zurich

Henri Corbaz, Director CHUV

Alexandre Oettli, Service cantonal de l'information statistique Vaud

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

In Switzerland, as in many other countries, risk adjustment is currently subject to much discussion. In fact, the new Federal Law on Sickness Insurance (KVG), in force since January 1st, 1996, introduced a retrospective risk adjustment scheme in order to reduce the incentive of sickness insurance funds to avoid bad risks and appeal to good risks (cream skimming). However, it is widely recognised that the risk adjustment mechanism currently in place is far from successful.[1]

One reason why cream skimming still occurs, and is a handicap for reducing costs through price competition among the sickness insurance funds, is that the present risk adjustment model is inadequate. It is based merely on age, gender and place of residence within each canton. In addition, risk adjustment is calculated retrospectively rather than prospectively, and may thus be considered as a "cost adjustment" scheme rather than a risk adjustment mechanism. Appropriate prospective risk adjustment models are considered to be an essential regulatory mechanism to ensure competition between sickness insurance funds.

The aim of this study is to develop a prospective health-based risk adjustment model for sickness insurance funds in Switzerland which, in addition to age and gender, uses medical information from prior use of health care services. In fact, because information on ambulatory diagnoses is currently not available in Switzerland, we will limit our presentation to risk adjustment models which use diagnostic information from prior hospitalisation.

At this stage we would like to emphasise that we have achieved our main objective, which is to demonstrate not only that it is possible to develop a health-based risk adjustment model in Switzerland using medical information from prior hospitalisation, but also that the type of risk adjustment procedure

[1]For a presentation and discussion of the risk-adjustment model in Switzerland see Beck (1998, 1999), Beck and Zweifel (1998), Spycher (1999), and Beck, Spycher, Holly and Gardiol (2001).

suggested in our study is much more effective than the model presently in force, which uses only demographic and geographic factors.

To understand the importance of the results presented in this study and their relevance to Switzerland, it may be useful to provide a brief description of the following aspects: the nature of risk selection and risk adjustment issues within the framework of the KVG; how we address these problems; and how we develop a solution.

1.1 Nature of the problem

At the outset, it is important to observe, as explained in Section 2, that establishing risk-related premiums is inherent to the efficient functioning of a health insurance system. If the risks are quantifiable, each individual pays the insurer an actuarial premium and receives, if he falls sick, an integral settlement. The actuarial premium is equal to the average of the health care costs of a group of individuals with the same "risk profile". The task of the sickness insurance fund then is to establish groups of individuals with the same risk profile, called an "actuarial category" and thus establish the same actuarial premium by pooling the risk associated with the occurrence of the (random) illnesses.

Also, because the very nature of the concept of insurance consists in covering an unpredictable random expenditure, insurers are forced to deny benefits for a number of years for a pre-existing condition, which reoccurs after admission to the fund, such as a severe and chronic pathology which will surely generate future expenditures.

However, establishing risk-based premiums generates inequity between high-risk and low-risk individuals. In addition, some individuals may not necessarily have the financial means at their disposal to purchase insurance within a risk-based premiums system. To solve these equity problems, in Switzerland like in many other countries, health insurance (or sickness insurance as it is called in Switzerland) is divided into "basic insurance" and "supplementary insurance". Although risk rating, in which the insurer charges grouped premiums based on its experience with the groups, and exclusion of pre-existing conditions are

allowed within the supplementary insurance system, they are not allowed within the “basic insurance” system. Instead, sickness insurance funds must accept without any reservations applicants from a given location, irrespective of age, gender, and health status, and charge all subscribers the same premium, i.e. a “community-rated” premium.

However, solving the equity problem through a community-rated premium creates incentives for sickness insurance funds to carry out a form of selection known as “risk selection”, or “cream skimming”. This selection may occur because community rating implies predictable profits on low-risk consumers and predictable losses on high-risk consumers. Consequently, sickness insurance funds have an incentive to avoid bad risks and appeal to good risks. Furthermore, this incentive is reinforced by the fact that health care costs are generally extremely skewed. In our data, for example, 5 percent of insured persons account for 50 percent of the spending, whereas 70 percent of insured persons account for 18 percent. Expenditures thus are very highly concentrated in relatively few individuals. Presented with such a concentration, it is tempting for a sickness insurance fund to accept the good risks and avoid the 5 percent of individuals that account for 50 percent of expenditure. Also, since an insured person’s health expenditures from the previous year are a relatively good predictor of the next year’s expenditures, sickness insurance funds can easily identify those who are likely to generate high costs in the future.

The primary reason for implementing risk adjustment is precisely to correct for risk selection problems. To this end, the identification of the individuals in which expenditures are very highly concentrated is crucial to the effectiveness of risk adjustment. The failure to have adequate risk adjustment creates a perverse set of selection incentives. Therefore, the most important criterion for evaluating risk adjustment models is the incentive to select good risks insofar as they reduce sickness insurance funds’ incentives to select good risks.

Devising a good risk adjustment model to reduce cream skimming consists in adjusting the premium by redistribution, in such a way that sickness insurance funds are paid a reasonable approximation of expected future costs. To this end, one uses factors, known as “risk adjusters”, which could help predict health care expenditures of policyholders, but which do not create perverse incentives. It is illuminating to observe that, in this way, the procedure which consists in

adjusting the premium by means of "risk-adjusted payments" associated with risk adjusters is an attempt to mimic actuarial premium computation, since risk-adjusted payments are an addition to the community-rated premium. This, in an algebraic sense (adding or subtracting), is equal to the average of the health care costs on the group of individuals having the same "risk profile" as determined by the risk adjusters.

We would like to call the reader's attention to the fact that, although health care expenditures in the previous year appears to be the best single predictor of an individual's future health care expenditures, *using it as a risk adjuster creates perverse incentives, as it tends to favour an increase in health care expenditures, and thus, should not be used for this purpose.*

The most important risk adjusters discussed in the literature on alternative risk adjustment models are the following: age and gender (or "demographic" adjusters); diagnosis-based risk adjustment; information from drug prescriptions; self-reported health information; mortality; disability and functional health status. For an extensive analysis of the advantages and disadvantages of these risk-adjusters, the reader is referred to Van de Ven and Ellis (2000).

Besides the demographic risk adjusters (age and gender), the remaining risk adjusters mentioned above are related to individual health status. It is clear from the concentration of expenditures that health status is an important factor in predicting variations in annual per-person health care expenditures in risk adjustment models. There are, in fact, different ways to develop health-based risk adjustment models, according to the nature of the variables that are used to describe individual health status. A description of some of the health-based risk adjustment models is given in Section 3 of this study.

In relation to the risk adjustment model presently in place in Switzerland, we would like to stress that age and gender are well known to be weak predictors of individual health care expenditures. Recall that under the KVG, risk adjustment is made at cantonal level. Within each canton, only age and gender are used as risk adjusters. The adult policyholder is classified according to a risk category. The first category consists of policyholders aged between 18 and 25 years. After that, policyholders are allocated to a category in 5-year brackets. The last group consists of policyholders aged 91 and older. Since gender is also a risk adjuster,

the risk categories number thirty in all. It is also worth noting that those policyholders aged 18 and less are not taken into account in the risk adjustment model. Each canton computes the risk adjustment as follows. First, the average value of costs of all the insurers in the database is determined. Second, the average value of costs in a given risk category is computed. The difference between these two averages indicates if sickness insurance funds receive a contribution[2] (if the sign is negative), or pay a contribution[3] (if the sign is positive). In this way, the sickness insurance funds must pay a (solidarity) contribution for the young policyholders and receive a "subsidy" for elderly policyholders.

Despite its merits, compared to a situation where no risk adjustment is performed, the present risk adjustment model does not fully provide incentives to prevent cream skimming. Indeed, apart from the fact that it is a retrospective model, it assumes that health care costs are only correlated with the two variables of "age" and "gender", and therefore does not properly take into account situations, such as young people with a very costly illness or, elderly people in good health. Therefore, by using risk adjusters related to health status, one should expect an important improvement on the present risk adjustment model which uses only age and gender.

Recently, Beck (1999) has suggested a model which takes into account individual health status. Beck (1999) observes that an insured person who has received inpatient treatment generates treatment costs that are seven times higher in the subsequent year than those who did not receive inpatient treatment. Based on this observation, Beck (1998, 1999) suggests a further two subdivisions of each of the 30 risk category in two sub-categories: whether the insured received inpatient treatment or not, thereby obtaining 60 risk categories. The main conclusion of Beck (1999) is that inclusion of prior hospitalisation as a risk adjuster has a strong impact on risk selection profits. Beck (1999) thus suggests the introduction of prior hospitalisation, in the form of a dichotomous variable ("yes" and "no"), as an additional risk adjuster to the formula currently in use. The suggested formula also changes the mode of calculation of risk adjustment from retrospective, as currently in use, to prospective.

---

[2] Or a "subsidy" in the terminology of van de Ven and Ellis (2000).

1.2 Our approach to health-based risk adjustment in Switzerland

The present study attempts to go beyond the important refinement of the existing risk adjustment model proposed by Beck (1999) and explores the possibility of introducing, at cantonal level, medical information from prior hospitalisation as a risk adjuster, alongside age and gender criteria.

In fact, we would have liked to develop a health-based risk adjustment model using diagnostic information from individual patient encounters in inpatient or outpatient services. Unfortunately, this is not feasible in Switzerland, as no reliable patient classification system for outpatient services is currently available. The results presented in this study must be viewed, however, as an important first step toward a full prospective risk adjustment model, incorporating both inpatient and outpatient care.

The development of a risk adjustment model based on inpatient encounter data requires a patient classification system (PCS) that ascribes each individual patient encounter to a group. In this study, we have decided to use the PCS known as "All-Patients Diagnosis Related" Groups (AP-DRGs, version 12), described in Section 5.2. Based on the patient's diagnoses, procedures, age and gender, the AP-DRGs assign each patient to a single, mutually exclusive group of patients that are expected to consume similar amounts and types of hospital resources.

Note that AP-DRGs classification is the only PCS which is compatible with routinely available hospital records in Switzerland at this time. The main applications for AP-DRGs in Switzerland are hospital budgeting, funding and reimbursement. However, as we demonstrate in our study, the AP-DRGs can be used as building blocks by sickness insurance funds for a higher level classification for risk adjustment.

This is achieved through the development of another smaller classification system, similar to the diagnosis-based risk adjustment models, which are founded on the idea of certain diagnoses predict health care expenditures. In fact, some inpatient encounter data provide information not only on the present

---

[3] Or a "solidarity contribution" in the terminology of van de Ven and Ellis (2000).

health status of the patient but also on its probable evolution. In particular, chronic or recurrent diseases will predict poor health status (high future expenditure for sickness insurance funds), while hospitalisation due to accidents will only have a short-term impact on individual health status. More specifically, based on a clinical approach and the main pathologies (Woolf, 1998), we have constructed a small patient classification system, independent of the AP-DRGs, consisting of 17 medical risk categories, which we called "medical adjustment categories".

The next step in the procedure consists in mapping the AP-DRGs in the medical adjustment categories. This is described in Appendix A. Consequently, we were able to construct a risk adjustment model based on the patient's diagnoses and procedures considered in the AP-DRGs, and classify each patient to a category of patients which takes into account the organs affected and the severity of the illness.

Our choice of AP-DRGs as the first step in the above procedure is due to the fact that this particular PCS is well known among those who are involved in discussions in Switzerland concerning hospital budgeting, funding and reimbursement according to sets of case types. In addition, some cantons have required public hospitals to implement invoicing systems based on AP-DRGs. We are perfectly aware of the limitations of the AP-DRGs as far as risk adjustment is concerned. First, they do not explicitly adjust for the severity of the illness, and this feature is somewhat at odds with the rationale behind a prospective risk adjustment model which seeks to predict health care expenditures (mostly relative to chronic illnesses) for the next year based on expenditures in the current year. Second, the AP-DRGs assignment rule is based on both patient diagnoses and procedures, whereas risk adjustment models should be mainly diagnosis-based and exclude, as far as possible, treatments or procedures, since classification is more difficult because the underlying illness is not known.

We intend, in our future work, to consider other types of PCS which explicitly take into account the severity of the illness and which are mainly diagnosis-based. Some of these PCS are described in Subsection 3.3.1. However, we would like to emphasise at this stage that we have achieved our main objective, which is to demonstrate, not only that it is possible to develop a health-based risk adjustment model in Switzerland using medical information from prior

hospitalisation records, but also that the type of risk adjustment procedure suggested in our study is much more effective than the model currently in use, which uses only demographic and geographic factors.

Meanwhile, without waiting until we consider these other types of PCS, we have examined a PCS known as SQLape, which is still in the development phase. The main application for which SQLape was originally developed was to adjust for health quality measurement indicators. However, in the same way as for AP-DRGs, we can map SQLape in medical adjustment categories. This is described in Appendix B. SQLape addresses one of the limitations of AP-DRGs; it does not take into account most secondary illnesses often linked to chronic illnesses, which are the best predictors for future health care expenditures. Here again, we do not claim that SQLape is the best alternative to AP-DRGs for the development of health-based risk adjustment. The purpose of using the SQLape in our study is to illustrate the potential improvement that could be achieved by using alternative classification systems which explicitly take into account the severity of the illness.

1.3 How did we proceed

The risk adjustment models presented in this study have been developed and estimated using data from people living in the canton of Vaud who had a health insurance policy with CSS, Helsana/Progres during the 1998–2001 time interval. They were validated with data from people living in the canton of Zurich, who also had a health insurance policy with the same insurance companies in 2001. We have been able to merge hospital records and insurance data into a unique database using a procedure which guarantees strict confidentiality. This procedure is described in Section 6.

The method for calculating the risk-adjusted payments to sickness insurance funds is outlined below. We considered linear regression models where the dependent variable represents expenditures, namely sickness insurance funds' payments to providers for both inpatient and outpatient services. Deductibles and co-payments for the covered services that were assumed by the policyholders were excluded from expenditures.

Our study examines two types of risk adjustment models, which we call the "one-year model" and the "multi-year model" respectively. The one-year model consists of predicting expenditures (i.e. sickness insurance funds payments to providers) in the current year (t) for the next year (t+1) using data from hospitalisation (gender, age, hospital and nursing home stays, medical adjustment categories) in year (t). As indicated earlier, medical adjustment categories are obtained alternatively from AP-DRGs and SQLape classifications.

In the same way as in Lamers and van Vliet (1996) and Lamers (1997), we also considered the multi-year risk adjustment model which, consists of predicting sickness insurance funds expenditures in the current year (t), for the next year (t+1) with available data from prior hospitalisation from previous years (t, t-1, t-2, etc. with a maximum of four years in our study). The multi-year models are particularly useful from a risk adjustment perspective, as most chronic illnesses can be cumulated over many years, i.e. the condition is considered to exist, if it occurred during any previous year. However, some conditions do not have a "multi-year" effect. To this end, we have constructed what we call a set of "aggregation rules" in order to better identify the risk adjustment groups on the basis of medical information from prior hospitalisation in previous years.

Although the aim of our study was to develop prospective health-based risk adjustment models, we carried out some intermediary work in order to judge the predictive power of a health-based risk adjustment model based on the medical adjustment categories suggested in this study. Specifically, as a preliminary step toward the development of a predictive risk adjustment model, we compared the one-year model with medical adjustment categories with the model currently in practice which only uses demographic information. In this comparison approach, we use the sickness insurance funds' expenditures observed in 2002 as a test sample. We estimate the models excluding year 2002 and subsequently predict year 2002. The predictions are then compared with actual expenditures.

On completion of this preliminary work, we developed one-year and multi-year prospective health-based risk adjustment models for the canton of Vaud[4], and compared them to other risk adjustment models using the same data. The

[4] The data for the canton of Zurich required for this part of the study were not available.

"demographic risk adjustment model", based only on age and gender, which we considered in our study, is very close to the model presently in force in Switzerland. Our results show that these health-based risk adjustment models dramatically improves upon the demographic risk adjustment model. They also significantly improve upon the type of risk adjustment model proposed by Beck (1998, 1999). These results highlight the great importance and the urgent need to establish an appropriate health-based risk adjustment model using medical information from prior hospitalisation.

Note, however, that similarly to the other prospective health-based risk adjustment models available in several countries, those which we developed underestimate high risks and overestimate low risks. This is also true, for instance, for the quite successful Principal Inpatient Diagnostic Cost Group (PIPDCG), which is the health-based risk adjustment model that the Health Care Financing Administration (HCFA) implemented in 2000 [Pope et al. (2002)]. Because of the underestimation of high risks, it would be useful at a later stage to combine a prospective risk adjustment model with retrospective risk sharing, in which sickness insurance funds are retrospectively partially reimbursed. This type of payment for sickness insurance funds was originally proposed by Ellis and McGuire (1986) and Newhouse (1986). The possible arrangement regarding retrospective payment of sickness insurance funds can include several forms of risk sharing, such as stop loss, or "risk sharing for high-risks" types of reinsurance mechanisms proposed by Van de Ven and Van Vliet (1992), and later analysed further by Van Barneveld et al. (1996). However, the present study does not address the problem of issue of combining a prospective risk adjustment model with a retrospective risk-sharing mechanism. This will be the subject of future research.

1.4 Plan of the study

This study is essentially divided in two parts. The first part consists of Sections 2 through 4. It contains a description of some of the main aspects of risk adjustment literature which we believe to be useful in assessing the interest and usefulness of our results. Our results are presented in the second part of this study (Section 5). The reader who is already well versed in all

aspects of risk adjustment models, including their estimation, and who is pressed for time, should go directly to Section 5.

The plan of this study is as follows. In Section 2, we gather some of the theoretical discussions on relevant risk adjustment and risk selection issues. Section 3 contains general considerations about risk adjustment. This includes clarification of the aims of risk adjustment models, a discussion on risk adjusters, a general discussion on some aspects of "patient classification systems" (PCS), and a brief presentation of possible arrangements by which sickness insurance funds could be paid partly on the basis of a prospective risk adjustment scheme, and partly retrospectively on the basis of the actual costs generated by their policyholders, notably "stop loss" or "risk sharing for high risks" reinsurance mechanisms. Section 4 briefly summarizes the theoretical aspects of the econometric approach used in this paper.

Section 5 describes the particular patient classification system we have developed for this study. It consists of a small number of categories, which we call "medical adjustment categories". This section also explains the mapping of the AP-DRGs and the SQLape into the medical adjustment categories. Section 6 describes the data and methodology used in this study. This section contains, among others, a detailed description of the studied population and of the matching procedure which was used to construct the data set. Some important intermediary econometric results are presented in Section 7, while the development of a prospective risk adjustment model and the corresponding econometric results are presented in Section 8. The implementation of the proposed prospective risk adjustment model is described in Section 9. A discussion on some topics which will be the subject of future research is provided in Section 10. Finally, the presentation of our main findings, and the feasibility of the risk adjustment model developed in our study are presented in Section 11.

1.5 Appendix: An illustrative example

We find it useful to awaken the reader's interest by describing some issues to be addressed in our research project by means of a very simple and illustrative example.

Consider a country where there are two classes of individuals: the “Good Risks” and the “Bad Risks”. Individuals in each group have different risks. We model that by stating that the patients may have three conditions, “low costs”, “high costs” and “severely ill”. Every year, each individual is randomly assigned to one of these three classes, independently of past conditions. The probabilities are shown in the following table where the currency in that country is represented by the symbol ☺:

Costs	Good Risks	Bad Risks	Total
Low cost	80%	0	**1** ☺
High cost	12.5%	2.5%	**5** ☺
Severely ill	2.5%	2.5%	**19** ☺
	95%	5%	

We assume there are 5% Bad Risks. The Bad Risks have a probability of 0.5 of being severely ill in the next year, and when they are not severely ill, the Bad Risks generate higher costs than the Good Risks.

Good Risks are less likely to be severely ill, and have a high probability of generating low costs. We chose our example in such a way that the Bad Risks represent 50% of the severely ill, and 1/6 of the high-cost group.

The average cost for the Good Risks is **2** ☺ [=(80*__1__ +12.5*__5__ +2.5*__19__)/95 ☺], and for the Bad Risks **12** ☺ [=(2.5*__5__ +2.5*__19__)/5 ☺] for a general mean cost of **2.5** ☺ [=(95*__2__ +5*__12__)/100 ☺].

The premium will therefore be 2.5 ☺ and the sickness insurance funds will gain 0.5 ☺ per Good Risk and lose 9.5 ☺ per Bad Risk.

It makes no difference to the sickness insurances funds if they gain 19 new Good Risks policyholders or if they get rid of one Bad Risk. They are also indifferent to a prevention mechanism lowering the costs of 95 individuals by 0.1 ☺ (10% of low costs) and getting rid of one Bad Risk. Obviously in this setting, the funds have a strong incentive to skim off the Bad Risks.

If we further assume a cost of 1 ☺ for getting rid of a policyholder, they will select out any group having an expected future cost above 3 ☺.

Assuming they do not collect any additional information, sickness insurance funds will select out anyone who was severely ill last year, as their mean cost is **7** ☺ [=(2.5*__2__ +2.5*__12__)/5 ☺], but also the high-cost group as their mean cost is **3.66** ☺ [= (12.5*__2__ +2.5*__12__)/15 ☺].

The impact of selection is not only on the 5% Bad Risks, but also on 20% of the population, since 15% of the population are actively selected out by the sickness insurance funds even though they are Good Risks, simply because they can not be identified as such.

One way to prevent this situation arising is through risk sharing. If the State sets up a pool which pays all costs generated by the severely ill, the average remaining costs for the Good Risks is 1.5 ☺ [=(80*__1__ +2.5*__5__ +2.5*__0__)/95 ☺] and for the Bad Risks **2.5** ☺ [=(2.5*__5__+2.5*__0__)/5 ☺]. This would mean that there would no longer be any selection since the gain from selection is exactly equal to the selection cost. However, this is expensive since the budget of this special fund would be 0.95 ☺ per individual. This solution is equivalent to adopting prior hospitalisation costs as a risk adjuster.

If the State establishes a procedure by which cost sharing only applies to one half of the costs generated by the severely ill, the sickness insurance funds will only select out severely ill individuals, and keep the high-cost group. The expected costs for the Good Risks are **1.75** ☺ [=(80*__1__+12.5*__5__+2.5*__9.5__)/95 ☺], and **7.25** ☺ for the Bad Risks. The expected costs for the high-cost group are then **2.66** ☺ [=(12.5*__1.75__ +2.5*__7.25__)/15 ☺], with the selection gain being less than 1 ☺.

Another approach attempts to identify the Bad Risks. This seems very difficult to implement. However, half of the severely ill will have had at least one inpatient stay during the previous year, making it possible to determine whether their health status is bad or good.

Let us assume that the medical record simply states that the patient's health status is good or bad. We may then adjust risks with respect to the conditions, instead of risk sharing when the patient is severely ill. The sickness insurance funds will receive 10 ☺ back for each of their policyholders known to be a Bad Risk. We do not identify all Bad Risks, only those who have been severely ill in the past. The incentive for the sickness insurance funds to select the severely ill group will cease to exist due to the distinction between Good and Bad Risks.

Even though the sickness insurance funds did not receive any transfers for the severely-ill Good Risks, they will not select them out, since they now know they are Good Risks. However, the situation remains unchanged for the high-cost group.

The real impact of this scheme will only be seen in the subsequent year. As half of the Bad Risks were detected during the previous year, the number of Good Risks in the high-cost group is now 10 times greater than the number of Bad Risks. Therefore, the expected cost for that group is only 32/11 or slightly less than 3 ☺. The difference in expected costs between them and the low-cost group is lower than 1 ☺. This should therefore put an end to the selection problem.

After a certain period of time, the 5% Bad Risks are all identified, and the risk adjustment pool fund only transfers 0.50 ☺ per individual.

This solution is much less expensive than cost sharing. (0.50 ☺ per individual instead of 0.95 ☺).

This simple model outlines the three important aspects we want to focus on in this research:

- Immediate identification of every bad risk is not the solution;
- The risk adjustment payment does not have to be computed for every patient; instead, the patients must be classified according to certain categories (here good and bad Risks), and risk adjustment payments only need to be computed for each category;
- At the social level, a major gain of the risk adjustment model is for individuals who are not risk adjusted. Namely, the 15% of individuals who were misclassified as potential bad risks.

2 RISK SELECTION AND RISK ADJUSTMENT ISSUES

In this section we gather some of the theoretical arguments which explain why, under some circumstances related to prevailing incentive mechanisms, the potential problem of "cream skimming", or "risk selection" by sickness insurance funds may exist. In turn, this problem has an indirect impact on costs, outcomes and quality of care. We also explain why well-designed risk adjustment schemes alleviate risk selection.

2.1 THE "FIRST-BEST" WORLD

At this stage, it is useful to recall the main features of an "ideal" model within the "paradigm of competitive economy", which would serve as a benchmark for the situation which actually prevails in most health care markets.

At the heart of this paradigm lie the first and second welfare theorems[5]: first, free markets are remarkably powerful at co-ordinating activities, and the corresponding allocation of scarce resources is Pareto efficient; second, equity issues may be solved with simple redistribution policies, such as lump sum transfers. However, this "first-best" vision of the world relies on many assumptions: consumers make rational choices, are perfectly informed about the consequences of their individual decisions, and internalise all costs and benefits of these decisions; competition between firms is perfect, all firms are small and do not collude, no co-ordination problems occur within the firm and production is efficient; there exists a complete set of markets for all conceivable goods. This generated a coherent body of knowledge, which remains extremely influential among economists.

[5] Arrow and Debreu (1954).

For future reference, we shall now recall some of the consequences of the "first-best" world for the health insurance market. First, competition between insurers must lead in theory to a better division of national consumption between health and other needs (this corresponds to the so-called "allocative efficiency"), and with a better "productive efficiency".

Among the assumptions of the first-best world, one supposes that, where there is uncertainty (the future health of the individuals is not known with certainty), the individuals have a risk aversion and insure themselves completely against the realisation of a random event which results in a loss ("a damage"). To insure oneself completely means that one buys an insurance policy so that one's wealth is unaffected in the event of damage occurring (if one falls sick) or in the event it does not occur.

If the risks are quantifiable, each individual pays the insurer an actuarial premium and receives, if he suffers damage (falls sick), an integral settlement.

The actuarial premium is equal to the average of the damage on the group of individuals with the same "risk profile"; the task of the insurer is to establish the groups, called "actuarial categories" and thus establish the same actuarial premium by pooling the risk associated with the occurrence of the (random) damage. Insurers offer as many contracts of complete insurance as there are actuarial categories. It is thus important to observe that *setting risk-related premiums is inherent in insurance.*

Let us now consider the implication of the above argument regarding allocative efficiency. In the situation outlined above, each individual pays for his own risk. Moreover, the allocation of resources to the insurance sector (and indirectly to the health care sector) reflects individual preferences.

The assumption is that all individuals do not have the same aversion to the same risks. In this case, according to the first-best paradigm, competition would make it possible to tend towards an ideal situation where the consumers, performing a rational trade-off between various possible allocations of their resources, would freely choose their level of hedging and would pay the right actuarial price to cover themselves, where necessary, against the health-related risks (individuals without health insurance cover). An allocative efficiency is thus

obtained since the individual utilities are maximised, and as a consequence social welfare is also maximised.

There is a further assumption that competition improves productive efficiency, as competitors may find it beneficial to minimise the production costs of the service which they provide.

In short, in the context of the first-best paradigm, the theory predicts that, if the individuals can choose their health insurance coverage freely and rationally, then the aggregation of these individual choices leads to a best-possible distribution of the resources between health care and other uses of the national income. This result is obtained because the paid insurance premiums will reflect the relative prices of the covered diseases compared to other consumption. One obtains these right actuarial prices because competition leads to risk-based premiums. Lastly, a competitive system enables each policyholder to be covered only for the risks he fears.

It should be noted at this stage that there are limits to this model within the first-best paradigm itself, so that competition among health insurers does not necessarily lead to an efficient allocation of resources. One important reason is because the market cannot insure health completely. This is because the very nature of the concept of insurance consists in covering an unpredictable random expenditure. In particular, a serious and chronic pathology will almost generate costs for the individual concerned. Therefore, covering this type of consumption in the long term is not, strictly speaking, the concern of an insurance mechanism insofar as the risk is known. This leads insurers to exclude covering the expenditure related to certain stated diseases (pre-existing conditions), as they do not see this as their job.

Another important limitation is because individuals do not necessarily have the financial means at their disposal to purchase the best insurance coverage. Therefore, one cannot consider the market efficient insofar as an individual cannot obtain the desired or desirable insurance coverage

2.2 EQUITY CONSIDERATIONS AND "SECOND-BEST" WORLD

The most important criticism made of the first-best competitive paradigm is

the inequity which it generates. There are in fact two main types of criticism. One concerns the equity between people with high risks and low risks, and the other concerns access to health insurance according to income.

Setting risk-based premiums generates inequity between the high-risk and low-risk individuals. It can be considered unfair if those who are high-risk pay more than those who are low-risk, insofar as they are largely not responsible for the level of this risk. Incidentally, this poses the problem of the responsibility of an individual for his own health, and is a source of heated debate in some countries.

Also, within a risk-based premiums' system, the budget share devoted to health insurance is appreciably different for poor and rich individuals, and can even prevent access to health insurance, and even to health care for a certain population group.

To solve these problems of equity, a regulation mechanism is put in place in many countries, including Switzerland, in which the following conditions are generally imposed by the regulator:

a) The minimum basic health insurance is compulsory for all persons resident in the country. This prevents an extreme form of adverse selection, namely low-risk individuals not cross-subsidising the high-risk individuals by not buying health insurance coverage.

b) For the basic health insurance, sickness insurance funds must accept applicants without any reservations.

c) For the basic health insurance, a sickness insurance fund must charge the same premium to all its policyholders (i.e. a "community-rated" premium)

d) All residents have freedom of choice in terms of sickness insurance fund and period of coverage (open enrolment).

With these constraints it is clear that one moves away from the operating conditions of the first-best competitive paradigm. In particular, it becomes impossible to ensure allocative and productive efficiency.

Let us note that the inequity between high risks and low risks is eliminated, but the problem of inequity in terms of to income remains. This regulation of the

health insurance premium is regressive insofar as the insurance contribution made by the poor is the same as those of the rich.

In addition to equity considerations, it is well documented that the view of the world provided by the first-best competitive paradigm breaks down when informational issues are seriously taken into account.[6] Markets are no longer complete and fail to allocate resources in the efficient manner of the first-best competitive paradigm. Still, some allocations are more inefficient than others, but the "first best" efficiency concept of the Arrow-Debreu world needs to be replaced with constrained efficiency, or a "second best" analysis. Moreover, redistribution of resources may also alleviate some distortions: efficiency and equity issues cannot be disentangled as easily as the first-best competitive paradigm predicts.

2.3 "Risk selection" or "Cream Skimming" and Risk Adjustment

Community rating, introduced on the basis of equity considerations, implies that a sickness insurance fund must ask the same premium contribution from each individual, independent of the individual's additional risk characteristics. This creates incentives for sickness insurance funds to carry out a form of selection known as "risk selection", or "cream skimming". This selection may occur because community rating implies predictable profits on low-risk consumers and predictable losses on high-risk consumers. For example, the expected health care costs for someone with a chronic disease such as cancer plainly exceeds expected health care costs for an otherwise similar person with no chronic disease. As a result, sickness insurance funds have an incentive to accept low-cost individuals and shun high-cost individuals.

To reduce cream skimming, there is a strategy to devise a good risk adjustment model whereby the premium is adjusted, by redistribution, in such a way that sickness insurance funds are paid a reasonable approximation of the expected future costs of their policyholders. A perfect risk adjustment is one for which the adjusted premium for any risk category is equal to its actuarial premium. In this case, there is no selection. Of course, for a number of obvious

[6] See, for example, Stiglitz (1993).

reasons, it is impossible to set up such a perfect risk adjustment model. However, the objective of good risk adjustment is to minimise the incentive of the sickness insurance fund to risk select. Note, however, that risk selection may continue if an inadequate risk adjustment scheme is in place. This is because cream skimming refers to the actions of sickness insurance funds to exploit risk heterogeneity [Newhouse (1996)]. This is precisely the situation in Switzerland where the current risk adjustment model is inadequate. It only uses age, gender and canton of residence as risk adjusters, and sickness insurance funds thus select risks on the basis of individual health status.[7]

Note that the need for a risk adjustment mechanism is the consequence of community rating, introduced on the basis of equity considerations. In a recent study, Zweifel (2002) questioned community rating and suggested that a risk adjustment mechanism could be completely avoided by allowing insurers to establish risk-related premiums. They argued that the inequity between the high-risk and low-risk individuals could be compensated by allowing persons with modest financial circumstances to be entitled to state assistance with premiums. The KVG has already introduced this possibility, and a large number of persons already benefit from state assistance with premiums. If the suggestion of Zweifel (2002) is to be followed, then the financial risk of high-risk individuals will be entirely shifted to taxpayers rather than shared between policyholders and taxpayers, as is presently the case. Here, a larger number of persons will apply for state assistance with premiums. A situation where a large percentage of the population receives state assistance would be socially difficult to accept. For this reason, we do not find their argument very convincing and continue to explore the possibility of establishing risk adjustment mechanisms which would reduce cream skimming.

In general, the precise form of the risk selection that may occur depends on the additional information that sickness insurance funds have. Our purpose is not to describe in detail the ways by which “cream skimming” could be achieved, as the reader may find this description in the published literature (see, for example, Van de Ven and Ellis (2000)). Instead, we would like to emphasise the importance of reducing the incentive to select risks, and the fact that the selection incentive must be studied by means of the “second-best body of

[7] See Spycher (1999), Beck, Spycher, Holly and Gardiol (2001)

knowledge"[8] which has been progressively built up since the Seventies.

There are two main reasons for a serious attempt to reduce the risk selection incentive. The first relates to the fact that competition among sickness insurance funds will better serve consumers if it is channelled toward the efficient delivery of care to a given patient, rather than seeking policyholders who are good risks. As a consequence, although a reduction of the risk selection incentive is not designed *per se* as a cost containment measure, it is a necessary condition to encourage sickness insurance funds to play an active role in the efficient delivery of care. The second is due to the fact that, as a result of risk selection, a competitive insurance market may be (highly) unstable. This in turn reflects the incentive of consumers to seek low premiums.

2.3.1 The "Purchase of Care" function of health insurers

The primary objective of an adequate risk adjustment mechanism is to reduce the incentive to select risks. As we shall see, it also has the potential effect of encouraging health insurers to play an active role in the efficient delivery of care.

This result is a consequence of what may be called the "purchase of care" function of health insurers. Indeed, due to changes since the 1980s in the structures of the health insurance and medical care industries, it is widely accepted that health insurers have two functions nowadays. The first one is the traditional insurance function which is to reduce the volatility of the policyholder's income by pooling the risks of a large number of people and operating on the principle of the Law of Large Numbers. The second function of health insures may be called the "purchase of care" function. This is well recognised in countries where managed care organisations are well developed. However, it also exists in a weaker form within the unintegrated system of health care treatment.

Before describing more precisely the purchase of care function, it may be useful to recall the main features of a managed care organisation (MCO) (or an "organised delivery system"). Following Folland, Goodman and Stano (2001),

---

[8] Guesnerie (1995) surveys how the Walrasian paradigm of first-best economics was challenged by second-best studies.

MCO could be described as having the following characteristics:

- It is a network of organisations (for example, hospitals, physicians, clinics and nursing homes);
- It provides or arranges to provide a co-ordinated continuum (from primary health care to emergency surgery) of services to a defined population;
- This system is willing to be held clinically and fiscally accountable for the outcomes and the health status of the population served;
- It is tied together by its clinical (treatment) and fiscal (financial) accountability for the defined population;
- Most often the organised delivery system is defined by its association with an insurance product.

Also, according to Folland, Goodman and Stano (2001), analysts identify three particular mechanisms by which sickness insurance funds seek to contain costs and/or improve quality of care:

- *Selective contracting*, in which payers negotiate prices and contract selectively with local providers such as physicians and hospitals;
- *Steering* policyholders to the selected providers;
- *Utilisation review* of the appropriateness of provider practices. This utilisation review may be prospective (in advance), concurrent (at the same time), or retrospective (looking back).

More generally, the "purchase of care" function of health insurers consists of a set of instruments with which they seek to contain costs and/or improve quality of care. Even within the unintegrated system of health care treatment, health insurers now review utilisation. Therefore, utilisation review could be considered as the minimal form of a "purchase of care function". Selective contracting and steering are additional functions which distinguish managed care from the more standard fee-for-service care.

The purpose of exercising the purchase of care function by health insurers is to address the health care costs in a manner that would also appear to address criteria of economic efficiency.

Having described the purchase of care function of health insurers, we now turn to the description of the potential effect of adequate risk adjustment on cost containment and economic efficiency.

To this end, we observe that health care expenses are generally extremely skewed. In our data, 70 percent of policyholders account for 18 percent of spending, while 5 percent (respectively, 10 percent) of the policyholders account for 50 percent (respectively, 61 percent) of spending. Thus, there is a very high concentration of expenditures in relatively few individuals.

With such concentration and in the presence of inadequate risk adjustment, it is tempting for a sickness insurance fund to enrol the 70 percent of individuals that account for 18 percent of spending and avoid enrolling the 5 percent of the individuals that account for 50 percent of spending. Also, since policyholders' expenditures from the previous year are a relatively good predictor of the next year's expenditures, sickness insurance funds can easily identify those who are likely to generate high costs in the future. Therefore, the financial incentive toward dis-enrolment is quite clear. The failure to have adequate risk adjustment creates a perverse set of selection incentives. For, by doing so, a sickness insurance fund could reduce the premium to be paid for the health insurance it sells, and put itself in a better financial situation compared to its competitors in relation to premiums. Of course, sickness insurance funds have to monitor their risk selection procedure constantly as every consumer, whether a good or bad risk, has an incentive to want to pay a low premium. Although this may be an expensive procedure, one should expect that the gains obtained from risk selection are much higher than those obtained by improving the management of sickness insurance funds.[9]

Therefore, if one could reduce the incentive to select by designing an appropriate risk adjustment mechanism, then premium or quality competition among sickness insurance funds will put pressure on them to exercise their role

---

[9] In Switzerland, the administrative costs of sickness insurance funds are about 4.5 percent of the total premium payments.

of “purchasers of care” more actively. Through this function, it is expected that sickness insurance funds collaborate more closely with health care providers in order to seek more economically efficient treatment, while preserving, if not improving, the quality of care.

2.3.2 Competitive insurance market and the “Death Spiral” problem

To illustrate the “Death Spiral” problem, let us consider a very simple setting, with only two types of health status, “Good” and “Bad” and two insurers who set their premium according to the average expected expenditure.

If the insured are equally distributed and the premiums proposed by the two insurers are equal, nobody will switch to the competitor, and the situation next year will remain exactly the same. Therefore, the insurance market is at equilibrium.

However, this equilibrium is not stable. In fact, if due to some random shock or deliberate action, the risks are not equally distributed, the average spending of the insurer with a slightly higher proportion of Bad Risks will rise, and it will have to propose higher premiums than his competitor. As is commonly accepted, Bad Risks are less likely to switch to a competitor, so that this small price difference will induce a higher proportion of Good Risks to switch than Bad Risks.

This causes a greater difference in premiums between the two competitors, which will induce more Good Risks to switch and so on, until one insurer collapses. As a result, we observe that even a very small initial difference can lead to the collapse of an unlucky insurer.

Risk adjustment does stop the Death Spiral. Prior to setting his premium, the unlucky insurer knows that he will receive a correction for the disadvantageous case mix. If the adjusters are correctly defined, this correction leads him to set exactly the same premium as his competitors, the initial equilibrium premium.

The estimations of price elasticity of switching are rather small in Switzerland, but strictly negative and empirical evidence shows that switching is correlated to health status. Therefore, the present system in Switzerland does satisfy the theoretical assumptions of the Death Spiral model.

If no risk selection occurs in the market, and due to the small number of initial switching, the probability of a Death Spiral starting and its deadly conclusion is small.

Conversely, if an insurer does not play according to the rules, and actively selects out Bad Risks, the market would immediately fall out of equilibrium, going directly to a situation similar to the latter stages of the Death Spiral.

The observation of the relative case mix between insurers which is directly related to premium levels shows that Switzerland is either in the very late stages of a randomly initiated Death Spiral, or that some insurers are cream skimming.

With the introduction of appropriate risk adjusters, the gain from selecting patients drops immediately to zero, so that the market would go back to the initial equilibrium and regular competition between insurers may be restored.

One could be tempted to assume that active risk selection is part of the competition between insurers, and the market rewards the more competitive insurers who optimally select risks. This is not the case in the health insurance market; insurance is compulsory, so that selection does not induce any gains, but only transfers the burden of bad risks to a competitor.

Such competition would lead to the collapse of the “normal” insurer. The bad risks would then have to go back to the selective insurer. The final stage would attribute the whole market to the best “competitor”, that is the one which is good at detecting bad risks, not the one which is the best at keeping health costs down, or optimising administration costs.

The active selection of good risks, which occurs by means of advertisement or the inclusion of a “bonus” which benefits only the healthy, should not occur in regular competition in health insurance. Even if the problem is less acute than that of the dis-enrolment of bad risks, since the markets gives a bonus to healthy people instead of inducing unnecessary costs to bad risks, the gains in terms of competition and market performance are still poor, and they also give rise to a Death Spiral.

The health insurance market should focus mainly on bad risks, and give them the best possible treatment for the lowest possible cost, rather than give the better bonuses to good risks. In normal correctly risk-adjusted competition, there would be an incentive for insurers to provide good conditions to the

severely ill, since the insurers would receive the correct reward for doing so. Without adjusters, insurers will always try to provide the worst possible conditions to severely ill patients, so as to insure that they will switch to their competitor.

Introducing risk adjusters does not only enable risk selection to be offset, but also restores a situation in which insurers would compete according to those parameters which really matter.

3 RISK ADJUSTMENT: GENERAL CONSIDERATIONS

As noted in the previous section, for equity reasons sickness insurance funds charge a community-rated (i.e. same) premium to all their policyholders, independent of their additional risk characteristics. This risk pooling thus creates predictable losses for sickness insurance funds on their high-risk individuals, and thereby incentives for sickness insurance funds to select policyholders because they prefer low-risk over high-risk consumers ("risk selection" or "cream skimming"). The precise form of the selection depends on the additional information that sickness insurance funds have.

At this stage, it is useful to recall that, following Van de Ven and Ellis (2000), we use "risk adjustment" to mean the use of information on the insured person to calculate their expected health expenditures over a fixed interval of time (generally a year) and set subsidies for consumers or sickness insurance funds to improve efficiency and equity.

Dunn (1998) suggests making a distinction between two aspects of risk adjustment:

- The risks' estimate (modelling and calculation of the expected medical costs for individual categories of the insured population according to some criteria)
- The risk adjustment scheme itself (setting up a risk adjustment mechanism on the basis of the risks' estimate)

The risks' estimate is in fact the essential aspect since it implies a modelling of health care expenditure. Once a satisfactory model is available, establishing a risk adjustment mechanism becomes more a political than an economic decision. Several types of mechanisms can be set up, but overall it boils down to a redistribution of resources on the basis of the risks' estimate; the sickness insurance funds are paid or compensated according to the risk structure of their policyholders, taking into account the criteria retained in the risk adjustment model.

The estimate of risks, which is represented by the risk adjustment model, seeks to explain the annual health care expenditure of an individual according to various risk factors, often called risk adjusters. Therefore, this usually means a panel data model, in which the risk adjusters are nothing other than the explanatory variables of the risk adjustment model.

3.1 OBJECTIVES

Correcting for risk selection problems is the primary reason for implementing risk adjustment. Thus, the most important criterion for evaluating risk adjustment models is the extent to which they reduce sickness insurance funds' incentives to select good risks. In fact, the very high concentration of expenditures in very few individuals makes the identification of these individuals crucial to the effectiveness of risk adjustment.

Note, however, that the predictive accuracy of different models which are used to estimate expected health expenditures of policyholders is by far the most common criterion on which risk adjustment models are compared. Yet the goal of risk adjustment is not accuracy *per se*, but rather improved incentives and equity [Van de Ven and Ellis (2000)].

It may well be the case that models with higher predictive power are not suitable for risk adjustment purposes. For example, as we shall see below, past health care expenditures are perfectly correlated with actual expenditures, and provide a high predictive power for models that include them among their explanatory variables. However, these models are not suitable for risk adjustment purposes.

A common measure of the predictive power of different risk adjustment models, but by no means the only one, is the conventional R^2, which measures the proportion of variance in individual expenditures explained by a set of risk adjusters.

Note that, as usual in any regression analysis, the R^2 should be interpreted with care. This is particularly important in the risk adjustment context since there are many reasons why R^2 can be misleading and difficult to compare, as emphasised in Van de Ven and Ellis (2000). In particular, they noted that "For the interpretation of R^2-values presented in the literature it is important that the R^2 (as well as the total variation) may depend on: (1) the type of services under analysis; (2) the (sub)population under analysis; (3) the variation in explanatory factors; (4) the level of medical technology; (5) the year of the data analysed, and (6) the length of the time period being predicted."

We emphasise once again that one of the most important criteria to use in selecting risk adjustment models is the incentives they create for risk selection. For this reason, as noted in Van de Ven and Ellis (2000), one of the most useful assessments that can be done is to compare actual and predicted expenditures for selected non-random groups of interest. For example, whenever possible, one should try to present an out-of-sample comparison of actual and predictive expenditures for several types of chronic diseases from a risk adjustment perspective.

3.2 Risk adjusters

The risk adjustment mechanism should account for predictable variations in annual per-person health care expenditures insofar as these are related to health status.

In many countries, considerable research has been conducted on alternative risk adjustment models, using a wide range of information. The reader will find in Van de Ven and Ellis (2000) a detailed discussion of these models in groups defined by the kind of data used for predicting health care expenditure: demographics only, prior year expenditures, diagnoses, information derived from prescription drugs, self-reported health and functional health status measures, mortality, and other types of information. Below, we briefly describe the risk adjusters that have been used in risk adjustment schemes to date.

Because age, gender and canton could be used to predict health care

spending, and because the information on age, gender and canton is independent of medical care, it appears attractive in terms of incentives to use these variables as risk adjusters. Currently in Switzerland, the risk adjustment procedure is based on demographic (age and gender) as well as geographic variables (cantons). However, various studies have shown that the most serious drawback of using age, gender and canton as risk adjusters is that they are too crude and do not accurately reflect expected costs; they are weak predictors of individual health care expenditures.[10]

Because expenditures in one year are highly correlated with expenditure in the following year, one may be tempted to set up a risk adjustment scheme by regressing health care expenditures in a given year on prior year health care expenditures (together with other demographic and geographic variables). However, although prior year health care expenditures appear to be the best single predictor of an individual's future health care expenditures, *using it as a risk adjuster creates perverse incentives, and thus, should not be used as such.*

However, the risk adjustment mechanism can be improved by extending the set of risk adjusters with measures that are more directly related to health. Using information on an individual's health status could be achieved through different means, as we shall now briefly recall.

One could set up a risk adjustment scheme by using self-reported health measures derived from surveys. The most common type of measures relate to perceived health status along several dimensions, functional health status, and chronic conditions.

As noted by Van de Ven and Ellis (2000), the potential equity and inefficiency problem of inappropriate incentives related to prior utilisation as a risk adjuster may be reduced by combining prior utilisation with diagnostic information. In fact, as often emphasised in the literature, risk adjusters based on diagnostic information from prior health care utilisation are quite promising.[11] In particular, it has been shown that, in the context of capitation systems, extending a demographic model with inpatient diagnostic information improves the

---

[10] For a survey, see Van de Ven and Ellis (2000) and the references contained therein.

[11] See, for example, Epstein and Cumella (1988), and Giacomi, Luft, and Robinson (1995) for the capitation system

predictive accuracy of the risk adjustment model.[12]

3.3 PATIENT CLASSIFICATION SYSTEMS (PCS)

This section contains a general discussion on some aspects of "patient classification systems" (PCS) in relation to risk adjustment. To this end, we shall first present some general features of PCS, and then illustrate important aspects of risk adjustment models developed in recent years.

A "patient classification system" consists of a set of rules that ascribes each individual patient encounter, or "case" to a group. The patient population is then a mixture of cases, i.e. a "case mix", which is characterised by the proportions of cases in each group. The rules that ascribe each individual patient encounter to a group are determined so that the groups are as homogeneous as possible with respect to clinical criteria (e.g., diagnoses and procedures) and resource consumption. Clinical information is generally coded according to a diagnosis coding system, e.g. ICD-9, ICD-10 (WHO, 1997), and a procedure coding system, e.g. ICPM [WHO, (1976)], ICD-9-CM [NCHS, (2002)]. In Switzerland, the sole routinely available morbidity statistics are hospital medical statistics. In fact, as of 1st January 98, all Swiss hospitals are legally obliged to code each case using ICD-10 for diagnosis and CH-OP (Swiss procedures codes, which are very close to the ICD-9-CM list) for procedures. Since, there are generally more than 20,000 diagnostic and procedure codes, a grouping software, called a "grouper", is used in routine applications to ascribe individual patient encounters to less numerous and more manageable groups.

The original PCS developed by Fetter in the late Sixties at Yale University [Fetter, (1980)] consists of groups called "Diagnosis Related Groups" or DRGs. As we shall illustrate in our study, such groups can be used on their own or as building blocks for a higher-level classification. Today, there are several dozen PCS. Note that some confusion may arise in relation to certain other PCS, since there is a tendency nowadays to use the term "Diagnosis Related Groups" as a generic term to designate the groups of which it they are made up.

---

[12] See, for example, Ellis and Ash (1995), Ellis, Pope, Iezonni et al. (1996), Lamers (1998).

Although several countries have developed country-specific patient classification systems (for example UK and France), the main approach to PCS development consists in adapting existing commercial systems to country peculiarities. This is notably the case for the various PCS made available by 3M Health Information Systems (AP-DRG, APR-DRG, IR-DRG). For a recent review of patient classification systems, see Dormont et al. (2003).

The main applications for PCS are hospital budgeting, funding and reimbursement [Rodrigues et al., (2002)]. However, more recently, due to the importance of risk adjustment in capitation systems of payments of health plans, some PCS are specifically developed for the purpose of risk-adjusted capitation payment. In fact, individual patient encounters contain diagnostic information used previously by either inpatient or outpatient services. Various studies have investigated the possibility of developing risk adjustment models using this diagnostic information. For a review, the reader is referred to Van de Ven and Ellis (2000).

The basic idea underlying all diagnosis-based risk adjustment models is the concept that certain diagnoses predict health care expenditures. For this reason, diagnosis-based models begin by identifying a subset of all diagnoses that predict current or subsequent year resource use. The models usually continue by grouping ICD-9 or ICD-10 codes into more aggregated groups based on clinical, cost, and incentive considerations. For example, based on individual diagnoses, procedures, age and gender, the DRGs (as well as the AP-DRGs) assign each patient to a single, mutually exclusive group of patients that are expected to consume similar amounts and types of hospital resources. Although this study is based on the AP-DRGs, detailed information about this particular PCS can be found in Section 6.2.

Recall, however, that the initial motivation for developing the DRG classification system was for hospital budgeting, funding and reimbursement. Since DRGs, as well as the AP-DRGs, are assigned after the services are rendered, it is fair to characterise them as “case-mix adjustment” since they explain past use of resources. However, as we shall demonstrate in our study, the AP-DRGs can be used as building blocks for a higher-level classification of risk adjustment for sickness insurance funds. As explained below, these blocks, or risk groups, can be used to predict future resource use, and must be assigned

before the services are rendered. More detailed information about AP-DRGs is given in Section 6.2.

To this end, in the current year (t), we use diagnostic information from hospitalisation in year t to predict health care expenditures for the next year (t+1). In addition, since the empirical analysis of this study is based on data for a 5-year period, from 1998 to 2002, we shall also explore, as in Lamers and van Vliet (1996) and Lamers (1997), the usefulness of incorporating multi-year diagnostic information from prior hospitalisation in risk adjustment models for Switzerland, instead of using only information from the previous year.

3.3.1 PCS explicitly taking into account severity of illness

A possible limitation of the use of AP-DRG for a risk adjustment scheme is the fact that they do not explicitly adjust for severity of illness. However, more advanced versions of DRGs have been developed that contain severity of illness subclasses within each DRG.[13]

There are a number of other PCS which have been specifically designed to take into account the severity of illness. This is notably the case of the Diagnostic Cost Group (DCG) risk adjustment models which were originally developed by Arlene Ash et al. (1989) and later significantly expanded [Ellis et al. (1996a, 1996b); Ash et al. (1998); Pope et al. (1998a, 1998b)], and the Clinical Risk Groups (CRGs) risk adjustment model developed by Averill et al. (1999).[14]

CRGs are risk groups that can be used as the basis of risk adjustment in a capitated payment system. The CRGs are a clinical model in which each individual is assigned to a single, mutually exclusive risk group which relates the historical demographic and clinical characteristics of the individual to the amount and type of health care resources that individual will consume in the future. In addition, the CRGs were developed to include explicit severity of illness subclasses that describe the extent and progression of an individual's condition. More specifically, as described in Averill et al. (1999), the CRG assignment for an individual is based on a hierarchically structured and detailed clinical approach

---

[13] See Averill et al.(1998)

[14] The development of CRGs was a joint effort between 3M Health Information Systems, Actuarial Sciences Associates, and the National Association of Children's Hospitals and related institutions.

which is based on a five-phase process. In this process, the identification of individuals with multiple interacting co-morbidity diseases and their associated severity of illness have been emphasised. Note, however, that in Phase I a disease profile and history of past medical interventions is created, and in Phase III, the assignment of the severity of each primary chronic disease is made for inpatient as well as outpatient sites. Despite its interest, CRGs are impossible to use in countries like Switzerland, where the required data for Phase I and Phase III are not available.

In fact, because information on ambulatory diagnoses is currently not available in Switzerland, we shall limit our presentation to risk adjustment models which use only diagnostic information from prior hospitalisation.

Although we shall not use a member of the family of the Diagnostic Cost Group (DCG) risk adjustment models, it is useful to outline some of their features for future reference. The DCG risk adjustment models were originally developed by Arlene Ash et al. (1989). The earliest versions are "single hierarchy" models and used only principal inpatient diagnoses. A recent series of studies has significantly expanded the original DCG framework [Ellis et al. (1996a, 1996b); Ash et al. (1998); Pope et al. (1998a, 1998b)]. Similarly to the CRGs, a fundamental change in the expanded framework is that instead of a single hierarchy used to rank all diagnostic groups, the recent models use information on multiple conditions, and impose hierarchies on diagnostic groups only when they are clinically related to each other. For a more complete description of these changes, we refer the reader to Van de Ven and Ellis (2000).

In relation to the work presented in our study, it is useful to summarise some of the features of the Principal Inpatient Diagnostic Cost Group (PIPDCG), which is the health-based risk adjustment model that the Health Care Financing Administration (HCFA) implemented in 2000. Our presentation is based on Pope et al. (2002).

Modelling begins by clustering diagnoses into 172 Principal Inpatient Diagnostic Groups (PIPDxGs). The next step consists of selecting diagnoses for inclusion in the payment model. In this step, 75 of the 172 PIPDxGs that may be minor, transitory or non-specific are excluded from the payment model. The PIPDCG model being a diagnosis-based model generally excludes treatments or procedures. However, an exception is made for chemotherapy as a principal diagnosis, because it identifies a group of very ill individuals with high expected costs. Another exception is made for human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS). For an HIV-related condition, HIV/AIDS may not be set as the principal diagnosis. For this reason, individuals are classified into the PIPDxG HIV/AIDS based on either a principal or a secondary diagnosis. The result of these first two steps is to produce a list of diagnoses which may generate increased payments.

The third step consists in using the DCG algorithm [Ash et al. (1989) and Ellis and Ash (1995)] to rank individuals and their diagnoses by their expected future costs. The result of the DCG algorithm consists in 16 PIPDCGs ranked in descending order of their expected future costs [see Pope et al. (2000), Table 4]. Individuals with multiple diagnoses in a given year are uniquely assigned to the single diagnosis that is most predictive of higher future expenditures, thus establishing a “single hierarchy”. Individual PIPDCG scores are included as categorical variables in linear regression models and used to predict future costs.

In relation to our study, it is useful to point out that Lamers and Van Vliet (1996) and Lamers (1997) expanded the DCG framework by considering multiple years of hospitalisation. The rationale for this is twofold. First, having had a serious hospitalisation in a given year might predictably give rise to above-average expenditures not only in the year directly following but also, to a diminishing degree, in the years thereafter (without necessarily resulting in a

new hospitalisation). Second, by giving higher premium subsidies to people who have been hospitalised for certain diagnoses during one of the previous years (instead of only during the last year), the probability increases that a health plan will receive an appropriate premium for its chronically ill policyholders. The results of Lamers and Van Vliet (1996) indicate that predictive accuracy improves when DCGs are incorporated over a longer period in the subsidy formula. For example, for the five percent of policyholders with the highest costs in year t-4 the predictable losses in year t decreased from 88 percent (demographic model) to 62 per cent (1-year DCG model) and to 43 per cent (3-year DCG model) of the predicted costs.

3.3.2 Ambulatory patient classification systems

The three most widely known classification systems are:

- the Ambulatory Care Group (ACG) system which has been developed by Jonathan Weiner and colleagues at Johns Hopkins [Weiner et al. (1991), (1996)];
- the Diagnostic Cost Group (DCG) family of models, which has been developed by Arlene Ash, Randall Ellis, Gregory Pope and colleagues at Boston University and Health Economics Research [Ash et al. (1989, 1998); Ellis et al. (1996a, 1996b); Pope et al. (1998a, 1998b, 1999)];
- the Disability Payment System (DPS) which has been developed by Richard Kronick, and Anthony Dreyfus, [Kronick, et al. (1996)].

Note that diagnosis-based risk adjustment models have also been developed by Hornbrook et al. (1991), Clark et al. (1995), and Carter et al. (1997).

It is important to emphasise at this stage that if the diagnosis-related data are often reliably coded in inpatient settings, they are unfortunately not always recorded in outpatient settings. When they are, their reliability is often questionable [Dunn et al. (1996)]. This explains why in countries like Switzerland, one cannot currently develop a risk adjustment mechanism based on outpatient diagnostic information.

It is fair to say, however, that not all individuals suffering from chronic conditions with related predictable costs are admitted to a hospital in any given year. Therefore, the risk adjustment mechanism can be further improved by extending the set of risk adjusters with measures that are more directly related to outpatient morbidity.

A promising step to improve a risk adjustment mechanism based on health status information is to use prescription drugs to infer the presence of chronic conditions. In fact, outpatient care diagnoses and information on chronic conditions deduced from the prior use of prescribed drugs have been shown to be good predictors of future health care expenditures.[15] Once more, the reader is referred to Van de Ven and Ellis (2000) and the references contained therein, for a detailed description of the use of information derived from prescription drugs in some risk adjustment schemes.

It is useful, for future reference, to mention that Lamers (1999) has recently developed an outpatient morbidity measure based on information about chronic conditions deduced from the prior use of prescribed drugs. She suggested that chronic conditions could be clustered into six "*Pharmacy Cost Groups"* (PCGs). It has also been noted by Lamers (1999) that, although the PCGs enhance predictive accuracy of future health care costs, possibility of gaming may arise within some health care systems. For example, drugs could be prescribed by health care providers in a situation where it would be more appropriate to change one's lifestyle. Lamers (1999) concluded that the usefulness of PCGs as risk adjusters may be restricted because of perverse incentives. However, in a very recent publication, Lamers and Vliet (2003) suggest the use of a "revised" PCG model in order to prevent manipulation of the risk adjustment mechanism and to overcome the problem of perverse incentives.

[15] See, among others, Weiner, Starfield, Steinwachs, Mumford (1991), Weiner, Dobson, Maxwell, Coleman, Starfield, Anderson (1996), Clark, Korff, Saunders, Baluch, Simon (1995), Lamers (1999).

Because of the high health care expenditures prior to death, mortality has been suggested as an additional risk adjuster. However, some authors have criticised its usefulness for a number of reasons. For instance, Van Vliet and Lamers (1998) observe that most of the excess costs associated with the high costs of dying are unpredictable and conclude that mortality should not be used as a risk adjuster. Another argument has been put forward which questions the social acceptability, and possibly a perverse incentive, for a sickness insurance fund to receive a high subsidy in the event of a policyholder's death.

Finally, because disability and functional health status have been shown to be relatively good predictors of future expenditures [Thomas and Lichtenstein (1986); Hornbrook and Goodman (1996)], indicators of functional health status could be used as risk adjusters.[16] In fact, Newhouse (1986) considered disability to be an almost ideal risk adjuster, and some countries use an indicator of disability in their risk adjustment scheme. This is notably the case for Belgium, Germany and the Netherlands.

3.3.3 Some advantages and disadvantages of risk adjustment models which are based solely on inpatient encounter data.

Although, from a practical point of view, diagnosis-based risk adjustment in Switzerland is presently feasible for inpatient data only, it is useful to reflect on the strengths and weaknesses of a risk adjustment model based solely on medical information from prior hospitalisation. Inpatient admission represents a significant expenditure for sickness insurance funds, and is often a proxy for the severity of an illness. Indeed, individuals who are hospitalised for the treatment of serious illnesses in year 1 are expected to have higher expenditures in year 2 than individuals who are not hospitalised or who are hospitalised for less serious illnesses.

It is important to recognize that the primary disadvantage of inpatient-based risk adjustment systems is the distorted incentives for sickness insurance funds to select health care providers. Sickness insurance funds obtain higher risk-adjusted payments only if their policyholders are admitted to hospital. This

[16] Indicators of functional health status reflect someone's ability to perform various activities of daily living and the degree of infirmity.

incentive is contrary to the aim to encourage health insurers to play an active role in the efficient delivery of care, by exercising the "purchase of care" function that we described earlier. This is because sickness insurance funds can be penalised when they avoid unnecessary hospital admissions through a prospective or concurrent utilisation review of the appropriateness of provider practices. It is thus essential to mitigate this kind of perverse incentive for hospital admission with other aspects of developing a risk adjustment model. For instance, one could exclude, like in the PIPDCG model (see Pope et al. 2000), diagnoses that may be minor, transitory and non-specific.[17] Note, however, that this does not mean that the costs associated with these diagnoses are dropped from the risk adjustment payment; the risk adjustment model captures these costs by means of the demographic and geographic risk adjusters.

Also, like in the PIPDCG model, the risk adjustment model could be fully hierarchical and assign individuals to the single diagnostic category that is most predictive of higher future expenditures. In this way, the risk adjustment model does not reward multiple hospitalisations, since readmission for the same diagnosis, or re-hospitalisation for diagnoses which are predictive of lower or equal future expenditures does not affect the category assignment.

There is also a rationale for risk adjustment which seeks to exclude diagnoses from short hospital stays.[18] Excluding short-day diagnoses may also exclude diagnoses that are appropriately performed in an outpatient setting.

However, excluding short-day diagnoses also has the disadvantage of reducing the predictive power of the risk adjustment model. Therefore, one faces a trade-off between limiting the incentive to select and reducing the predictive power of the risk adjustment model in deciding whether to include or exclude short-stay diagnoses. Future expenditures associated with short-stay diagnoses are significantly lower than the future expenditures associated with long-stay diagnoses. Therefore, excluding short-stay diagnoses does not greatly reduce the predictive power of the risk adjustment model, and this could favour the decision

---

[17] As in the PIPDCG model, these could consist of the following diagnoses: ectopic pregnancy, miscarriage/terminated pregnancy, completed pregnancy with major complications, completed pregnancy with complications, completed pregnancy without complications (normal delivery).

[18] For example, PIPDCG model defines short-stays admissions as zero-day stays (i.e. the same admission and discharge dates) or one-day overnight stays (i.e. discharge date one day later than admission day)

of excluding short-stay diagnoses.

3.4 Risk sharing, Stop Loss and Reinsurance

As noted by Newhouse (1998), "It is far easier to describe the potential usefulness of risk adjustment than to do it well". In fact, experience shows that, in general, diagnosis-based risk adjustment models substantially under-predict expenditures for all chronic disease groups of policyholders, or for the top 1 or 5, or 10 percent of future expenditures. Indeed, no PCS can accurately classify 100 percent of patients. Some patients have clinical conditions so unique that they defy classification.

This observation implies that, even if we considerably improve the risk adjustment mechanism by better use of the diagnostic information, we may not be able to derive a risk adjustment mechanism that defines a sufficiently accurate lump sum for well-defined groups of individuals to prevent risk selection completely. For this reason it is essential to introduce other components of the payment system of sickness insurance funds. Following Newhouse and his co-authors (Newhouse 1986, 1991, 1994, 1996; Newhouse, et al., 1989, 1997; Keeler, et al., 1998), it could be suggested that sickness insurance funds be paid partly on the basis of a prospective risk adjustment scheme, and partly retrospectively on the basis of the actual cost of their policyholders. This suggestion corresponds to what is generally termed "supply-side cost sharing" in economics literature (Ellis and McGuire (1986), (1993); Selden (1990)).

At this stage, it is worth analysing risk adjustment models in the context of the medical care triad: sickness insurance funds, providers and patients. The government is supposed to act both as a regulator, which sets risk adjusters as prices, and as a "sponsor" which redistributes financial revenues among providers and insurers. Although risk adjustment is only one part of the payment system, the development of the classification system on which it is based cannot be carried out in isolation from the other components of the full payment system of sickness insurance funds and providers.

Since premiums cannot be perfectly risk-adjusted, the adverse effects of

selection may also be reduced by various forms of *ex post* risk sharing between the sponsor and sickness insurance funds, whereby the latter are retrospectively and partially reimbursed by the sponsor for some of their costs.

More precisely, in addition to the risk adjustment scheme, the payment of sickness insurance funds can include several forms of risk sharing, such as stop loss, or "risk sharing for high risks" types of reinsurance mechanisms proposed by Van de Ven and Van Vliet (1992), and later analysed further by Van Barneveld et al. (1996). The goal of risk sharing is to reduce the sickness insurance funds' *predictable* losses and profits, while preserving their incentives for efficiency as far as possible.

Note that the "risk sharing for high risks" completely differs from traditional reinsurance. Traditional reinsurance does not reduce the sickness insurance funds' predictable losses on high-risk individuals, and therefore, cannot be a tool to reduce the sickness insurance funds' incentives to select risks.

Under a stop loss system, all expenditures above a specific threshold are paid under the stop loss policy, which is separate from the risk adjustment mechanism. This means that all expenditures for an individual that exceed the stop loss threshold are excluded from the risk adjustment scheme. More specifically, if expenditures exceeded the threshold for each individual, then the expenditure amount would be set equal to the threshold, which means that the expenditures for each individual are capped at the stop loss threshold. Using capped expenditures, the risk adjustment would be used to compute risk-adjusted premiums.

Whereas in a stop loss model, the amount of expenditures covered by the risk adjustment mechanism is reduced, in a risk-sharing reinsurance model, this amount is not reduced, but the amount of the risk-adjusted premiums is increased for individuals with high expenditures. It is important to note that a reinsurance model is similar to an outlier policy, and in fact, one cannot dissociate the reinsurance payment scheme with the outlier policy used in the specific PCS-based payment of hospital systems. This outlier policy is considered as a cost-sharing mechanism between hospitals and sickness insurance funds.

In the present study, we shall not pursue the analysis of a "risk sharing for high risk" type of reinsurance, although we believe it would be worthwhile to conduct a detailed comparison of the effect of a reduction in the level of risk

associated with a given risk adjustment system and that of stop loss models. For a detailed analysis of reinsurance as a type of risk-sharing mechanism between the sponsor and sickness insurance funds, the reader is referred to Van de Ven and Ellis (2000) and the references contained therein.

4 ECONOMETRIC MODELLING

The purpose of this section is to summarize the econometric approach used in this paper. The material in this section does not constitute a survey of the literature on the econometrics of risk adjustment; this would be beyond the scope of this study.

4.1 GENERALIZED TOBIT

To this end, it is illuminating to first consider the situation where we want to model both the probability of using health care services and expenditure levels, conditional on a positive use. To this end, it is useful to introduce two latent variables y_{n1}^* and y_{n2}^*, where y_{n1}^* represents the propensity to use health care, and y_{n2}^* represents the propensity for the amount spent on health care.

We assume that y_{n1}^* and y_{n2}^* are explained by the following linear regression models:

$$\begin{aligned} y_{n1}^* &= x_{n1}^{'}\beta_1^0 + u_{n1}^0 \\ y_{n2}^* &= x_{n2}^{'}\beta_2^0 + u_{n2}^0 \\ \mathcal{D}\begin{pmatrix} u_{n1}^0 \\ u_{n2}^0 \end{pmatrix} &= \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^{02} & \sigma_{12}^0 \\ \sigma_{12}^0 & \sigma_2^{02} \end{pmatrix}\right) \end{aligned} \tag{1}$$

We further assume that y_{n2}^* is not observable, whereas y_{n1}^* is observable only when the propensity to use health care y_{n2}^* is positive.

Let y_{n1} denote the observable variable, which is formally defined by:

$$y_{n1} = \begin{cases} y_{n1}^* & \text{if } y_{n2}^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

The model specified by equations (1) and (2) is known in econometrics literature as Generalized Tobit.

Let us denote the set of individuals which have a zero expenditure as $\mathcal{I}_0$ (i.e. $y_{n1} = 0$) and the set of individuals who have a positive expenditure as $\mathcal{I}_1$ (i.e. $y^*_{n2} \geq 0$). We may write formally:

$$\mathcal{I}_0 = \left\{ n : y^*_{n2} < 0 \right\} \text{ and } \mathcal{I}_1 = \left\{ n : y^*_{n2} \geq 0 \right\} \quad n = 1, \ldots, N$$

It can be shown [see, for example, Gouriéroux (1989)] that the Likelihood function of this model is equal to:

$$\begin{aligned} \mathcal{L}_N = & \prod_{\mathcal{I}_0} \Phi\left(\frac{-x'_{n2}\beta_2^0}{\sigma_2^0} \right) \prod_{\mathcal{I}_1} \frac{1}{\sigma_1^0} \varphi\left(\frac{y_{n1} - x'_{n1}\beta_1^0}{\sigma_1^0} \right) \\ & \prod_{\mathcal{I}_1} \Phi\left\{ \frac{1}{\sqrt{1-\rho^{02}}} \left[\frac{x'_{n2}\beta_2^0}{\sigma_2^0} + \frac{\rho^0}{\sigma_1^0}\left(y_{n1} - x'_{n1}\beta_1^0 \right) \right] \right\} \end{aligned} \tag{3}$$

It is clear from (3) that the identifiable parameters are: β_1^0, σ_1^0, $\frac{\beta_2^0}{\sigma_2^0}$ and ρ^0

Instead of estimating the Generalized Tobit model by Maximum Likelihood methods, one can use a two-stage estimation procedure of the type suggested by Heckman (1976). The basic idea behind the two-stage method is to observe that for the nonzero observations y_{n1}, we have

$$E\left(y_{n1} \mid y^*_{n2} \geq 0 \right) = x'_{n1}\beta_1^0 + \rho^0\sigma_1^0 \frac{\varphi\left(x'_{n2}\beta_2^0 / \sigma_2^0 \right)}{\Phi\left(x'_{n2}\beta_2^0 / \sigma_2^0 \right)}, \tag{4}$$

where $\varphi\left(x'_{n2}\beta_2^0 / \sigma_2^0 \right)$ and $\Phi\left(x'_{n2}\beta_2^0 / \sigma_2^0 \right)$ are the density function and the distribution function of the standard normal evaluated at $x'_{n2}\beta_2^0 / \sigma_2^0$. Thus, equation (4) can be written as:

$$y_{n1} \mid y^*_{n2} \geq 0 = x'_{n1}\beta_1^0 + \rho^0\sigma_1^0 \frac{\varphi\left(x'_{n2}\beta_2^0 / \sigma_2^0 \right)}{\Phi\left(x'_{n2}\beta_2^0 / \sigma_2^0 \right)} + \varepsilon_{n1}^0 \tag{5}$$

where $E(\varepsilon_{n1}^0) = 0$.

The estimation of (5) by a two-stage procedure is the following: Define a dummy variable

y_{n2}=1 if $y^*_{n2} \geq 0$
y_{n2}=0 otherwise

Then, using the Probit model, we arrive at the ML estimates of β_2^0/σ_2^0. Consequently, we arrive at the estimated values of $\varphi\left(x_{n2}'\beta_2^0/\sigma_2^0\right)$ and $\Phi\left(x_{n2}'\beta_2^0/\sigma_2^0\right)$. Now, using (5) and the expression for $V\left(y_{n1} \mid y_{n2}^* \geq 0\right)$ we get consistent estimates of β_1^0, σ_1^0 and ρ^0.

The estimation of the expected future consumption is then equal to:

$$\begin{aligned} E\left(y_{n1}\right) &= P\left(y_{n2}^* \geq 0\right) E\left(y_{n1} \mid y_{n2}^* \geq 0\right) \\ &= \left(\Phi\left(-x_{n2}'\beta_2^0/\sigma_2^0\right)\right)\left(x_{n1}'\beta_1^0 + \rho^0\sigma_1^0 \frac{\varphi\left(x_{n2}'\beta_2^0/\sigma_2^0\right)}{\Phi\left(x_{n2}'\beta_2^0/\sigma_2^0\right)}\right) \end{aligned} \tag{6}$$

where the first term comes from the Probit regression, and the second from the OLS regression, including the Mills ratio

In the particular case where the disturbance terms u_{n1}^0 and u_{n2}^0 are uncorrelated (i.e. $\rho^0 = 0$), the Likelihood function (3) may be split in two parts, the first one identical to a Probit regression and the second is identical to the Likelihood of the OLS on the sub-sample of positive observations:

$$\mathcal{L}_N = \left(\prod_{\mathcal{I}_0} \Phi\left(\frac{-x_{n2}'\beta_2^0}{\sigma_2^0}\right) \prod_{\mathcal{I}_1} \Phi\left(\frac{x_{n2}'\beta_2^0}{\sigma_2^0}\right)\right) \prod_{\mathcal{I}_1} \frac{1}{\sigma_1^0} \varphi\left(\frac{y_{n1} - x_{n1}'\beta_1^0}{\sigma_1^0}\right) \tag{7}$$

Thus, in this case the Heckman procedure simplifies it in a two-step procedure, first a Probit equation to compute the probabilities of positive expenses, and the second step is a regular OLS estimated on the sub-sample of positive expenses.

$$\begin{aligned} E\left(y_{n1}\right) &= P\left(y_{n2}^* \geq 0\right) E\left(y_{n1} \mid y_{n2}^* \geq 0\right) \\ &= \left(\Phi\left(-x_{n2}'\beta_2^0/\sigma_2^0\right)\right)\left(x_{n1}'\beta_1^0\right) \end{aligned} \tag{8}$$

4.2 TRANSFORMING THE DEPENDENT VARIABLE

At this stage, it is useful to recall the three main characteristics of health care expenditures:

a A large percentage of the population has no medical care expenditure during any given year.
b The remaining percentage has positive expenditures that are highly skewed. Through much of their range, the positive expenditures are approximately log-normally distributed.
c The far-right tail of the distribution is too long even for a log-normal distribution. The thickness of the tail stems from a small number of heavy users.

The Generalized Tobit model described above takes care of the first of these characteristics. Because of the second characteristics, some suggest transforming the dependent variable in order to achieve (approximate) normality. Here some authors propose replacing y_{n1} by $\ln(y_{n1})$ [see, for example, van De Ven and van Praag (1981)]

Using a different approach, Duan, Manning, Morris, Newhouse (1983) introduced the "Two-Part Model" (TPM). Although the relation between TMP and the Generalized Tobit has been the subject of heated debate, one could argue that the TPM model corresponds to the particular case of the Generalized Tobit model described above, where it is assumed that independence between u_{n1}^0 and u_{n2}^0 $\left(\rho^0 = 0\right)$, and the dependent variable is the log of health care expenditures.

In this case, the Likelihood function may be written as:

$$\mathcal{L}_N = \prod_{\mathcal{I}_0} \Phi\left(\frac{-x_{n2}'\beta_2^0}{\sigma_2^0}\right) \prod_{\mathcal{I}_1} \Phi\left(\frac{x_{n2}'\beta_2^0}{\sigma_2^0}\right)$$
$$\prod_{\mathcal{I}_1} \frac{1}{\sigma_1^0} \varphi\left(\frac{\ln(y_{n1}) - x_{n1}'\beta_1^0}{\sigma_1^0}\right)$$

Although we certainly recognise the problems that may arise by using the untransformed dependent variable, we generally prefer to estimate the untransformed model. There are two reasons for this. First, recent literature has shown that the transformed models introduced additional problems which are difficult to solve [see Mullahy (1998), Manning and Mullahy (2001)]. Second, risk adjustment payments are made in monetary units and not in log-transformed monetary units, and their actual computation may be easier to understand, if it is done without transforming the dependent variable.

This does not mean that we ignore the problems which occur in estimating a linear model (untransformed) by OLS, given the skewness and kurtosis of the expenditure data. These problems are addressed in Holly and Pentsak (2003) [see also Manning, Basu and Mullahy (2003)]. However, in this study we shall not apply more elaborate procedures than Ordinary Least Squares (OLS) to the (untransformed) linear model. One reason being that in the type of PCS used in our study – as well as in many other studies – the untransformed linear model directly corresponds to the cell-based approach that is the form in which most risk adjusters are calculated, as we shall explain below.

4.2.1 Special case: mutually exclusive and exhaustive risk categories.

In many patient classification systems, individuals are assigned to mutually exclusive and exhaustive risk categories. This is notably the case of AP-DRGs on which one of the patient classification systems we use in our study is based. Moreover, the useful result from the procedure is the mean future costs for each of these risk categories, which corresponds to the risk-adjusted payment.

The results presented above implicitly assumed that the components of the vector of exogenous variables x_{n2} did not only consist of dichotomous (or indicator) variables. Therefore, the results presented above may be extended.

Mutually exclusive indicator variables have the following structure:

$$x_{n2}^{(k)} = \begin{cases} 1 & \text{if patient is in condition } k \\ 0 & \text{otherwise} \end{cases}$$

So that

$$X_2' X_2 = \begin{pmatrix} N_1 & 0 & \cdots & 0 \\ 0 & N_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & N_K \end{pmatrix} \; ; \; X_2 = \begin{pmatrix} \vdots \\ x_{n2}' \\ \vdots \end{pmatrix} = \begin{pmatrix} \cdots & x_2^{(k)} & \cdots \end{pmatrix}$$

Where N_k is the number of individuals in risk category k.

We define $y_2^{(k)} = x_2^{(k)\prime} y_2$ equal to the number of individual with positive expenditures in group k.

In this case, for members of the risk category k, $x_{n2}' \beta_2 = \beta_2^{(k)}$ since only $x^{(k)}$ is different from zero. The Likelihood for the Probit regression splits into k factors. Each factor is identical to the Probit likelihood function in sub-sample k.

The first step of the procedure will therefore give the following coefficient and probability:

$$\tilde{\beta}_2^{(k)} = \Phi^{-1}\left(\frac{N_k - y_2^{(k)}}{N_k}\right) \Rightarrow P\left(y_2 > 0 \mid x_2^{(k)} = 1\right) = \frac{y_2^{(k)}}{N_k}$$

The Mills ratio computed for this cell is a constant equal to $\dfrac{\varphi\left(\widetilde{\beta}_2^{(k)}\right)}{\Phi\left(\widetilde{\beta}_2^{(k)}\right)}$

In the case of mutually exclusive categories, the Mill's ratio is perfectly collinear with the X_2 matrix. The second step cannot be performed without imposing an additional identifying constraint. Two possible constraints are available: either consider two different mutually exclusive risk categories in the first and second step, or consider the two regressions as independent, that is $\rho^0 = 0$.

In our model, it would be impossible to construct different mutually exclusive risk categories for each step, so that we have to consider the two regressions as independent.

In this case, the second step is a regular OLS computed on the sub-sample with positive costs:

$$X_1^{'}X_1 = \begin{pmatrix} \ddots & & 0 \\ & y_2^{(k)} & \\ 0 & & \ddots \end{pmatrix}, X_1^{'}y_1 = \begin{pmatrix} \vdots \\ \sum_{\{i:y_{i2}=1\}} y_{i1} \\ \vdots \end{pmatrix}$$

In order to compute the risk adjuster for a particular risk category, we neither need the conditional mean nor the probability, but the unconditional mean. We want to reimburse two risk groups exactly in the same way, one with 50 with probability 1 and the second with probability .5 of expenditures of 100.

Therefore, the risk adjuster for any risk category is equal to:

$$\begin{aligned} \tilde{\beta}^{(k)} &= P(y_2 > 0 \mid x_2^{(k)} = 1) E(y_1 \mid y_2 = 1; x_2^{(k)} = 1) \\ &= \frac{y_2^{(k)}}{N_k} \frac{\sum_{\{i:y_{i2}=1;x_{i2}^{(k)}=1\}} y_{i1}}{y_2^{(k)}} = \frac{\sum_{\{i:y_{i2}=1;x_{i2}^{(k)}=1\}} y_{i1}}{N_k} \end{aligned}$$

The OLS does compute directly the unconditional mean, as we can verify below:

$$\begin{pmatrix} \vdots \\ \hat{\beta}^{(k)} \\ \vdots \end{pmatrix} = \left(X_1^{'}X_1\right)^{-1} X_1^{'}y_1 = \begin{pmatrix} \ddots & & 0 \\ & N_k & \\ 0 & & \ddots \end{pmatrix}^{-1} \begin{pmatrix} \vdots \\ \sum_{\{i:x_{i1}^{(k)}=1\}} y_{i1} \\ \vdots \end{pmatrix}$$

$$= \begin{pmatrix} \vdots \\ \frac{\sum_{\{i:x_{i1}^{(k)}=1;y_{i1}>0\}} y_{i1} + \sum_{\{i:x_{i1}^{(k)}=1;y_{i1}=0\}} 0}{N_k} \\ \vdots \end{pmatrix}$$

Noting that $y_{i1} > 0 \Leftrightarrow y_{i2} = 1$, we may conclude that $\tilde{\beta}^{(k)} = \hat{\beta}^{(k)}$. In other words, in this particular case, the OLS estimator and the two-step estimator are identical.

4.2.2 Non mutually exclusive risk categories

When the risk categories are not mutually exclusive, the estimators may diverge. To illustrate this point let us consider a case with three illnesses.

The whole population has a probability of p_0 of getting the first illness, with treatment costs C_0; only people with $x_1 = 1$ have a probability of p_1 to get the second illness with treatment costs C_1; finally, only people with $x_2 = 1$ have a positive probability of p_2 to get the third illness with treatment costs C_2.

The population may be grouped in four risk categories, labeled 00, 10, 01 and 11, according to the couples $\{x_1, x_2\}$.

For the four sub-samples, the distribution of expenses is:

risk category	$P(y=0)$	$E(y)$	$E(y \mid y>0)$
00	$1-p_0$	p_0C_0	C_0
10	$(1-p_0)(1-p_1)$	$p_0C_0 + p_1C$	$\frac{p_0C_0 + p_1C_1}{1-(1-p_0)(1-p_1)}$
01	$(1-p_0)(1-p_2)$	$p_0C_0 + p_2C_2$	$\frac{p_0C_0 + p_2C_2}{1-(1-p_0)(1-p_2)}$
11	$(1-p_0)(1-p_1)(1-p_2)$	$p_0C_0 + p_1C_1 + p_2C_2$	$\frac{p_0C_0 + p_1C_1 + p_2C_2}{1-(1-p_0)(1-p_1)(1-p_2)}$

We should first note that the OLS estimation would yield p_0C_0 for the constant term, p_1C_1 as the coefficient of x_1, and p_2C_2 as the coefficient of x_2. So that OLS coefficients are unbiased leading to an $R^2=1$

The first step of the two-step estimation would yield biased estimations of the probabilities in general, except for a particular density function for which

$$\begin{aligned} \Phi(\alpha) &= 1-p_0 \\ \Phi(\alpha+\beta_1) &= (1-p_0)(1-p_1) \\ \Phi(\alpha+\beta_2) &= (1-p_0)(1-p_2) \\ &\Downarrow \\ \Phi(\alpha+\beta_1+\beta_2) &= (1-p_0)(1-p_1)(1-p_2) \end{aligned}$$

In general, the first three equations determine completely the three

parameters α , β_1 , β_2 and the fourth equation is almost surely not satisfied. This condition is not satisfied with the normal distribution used in the Probit model, unless $p_1 p_2 = 0$, so that we defined unobserved conditions.

The two-step estimators are then also biased, since the Mill's ratio would not be of any help to correct for the first-step bias, since the model does not include any correlation between the two steps.

Thus, if the conditions are additive, that is if costs for treating illnesses 1 and 2 are both equal to the sum of the costs for treating illness 1 and illness 2 independently, the OLS estimators are better than the two-step estimator from a statistical point of view.

Moreover, we should note that the diagnostic groups are constructed in such a way that two diagnostics, one with 10% of future costs of CHF 1,000 and the second with 50% of future costs of CHF 200 could be grouped together in a group with CHF 100 as expected future costs. As we obviously see, however, these two conditions would not yield similar results in the first step (the coefficient of the first diagnosis should be smaller) and the second step (the coefficient of the first diagnosis should be higher). However, these two differences total zero when we look at the expected future costs.

The two-step procedure would be adequate if we could rely on two diagnostic groups, one to forecast the probability of positive expenditures, and a second to forecast the expected expenditures, once we know that the patient suffers from a disease.

For example, let us consider two diagnoses, one with 10% of future costs of CHF 1,000 and the second with 50% of future costs of CHF 200. The coefficient of the first diagnostic should be smaller in the first step, and higher in the second. These two differences total zero when we look at the expected future costs.

On the other hand, if we rely on OLS estimations, these two diagnoses may be directly grouped in the expected average costs of the CHF 100 risk category. This implies in turn that the risk category construction is simpler and easier to validate with a one-step procedure than in the two-step procedure, where errors in the construction of first-step risk categories would be accumulated with the second-step errors.

Therefore, in this study we shall only focus on risk category definitions which are homogeneous with respect to expected future costs, without looking closely at the first- and second-step adequacy of the groupers.

5 THE PATIENT CLASSIFICATION SYSTEMS USED IN THIS STUDY

5.1 GENERAL CONSIDERATIONS

As mentioned in Section 3.3, the aim of patient classification systems is to summarise the large quantity of data available, and group them in adequately defined classes to predict different outcomes, such as hospital expenditures, mortality rates or insurance risks. In the context of this study, the objective is to reduce the incentive for sickness insurance funds to select risks. It is important therefore to consider that the objective is not to predict current sickness insurance funds' expenditures, but future sickness insurance funds' expenditures. In this respect, medical information is interesting, because it provides information not only on the present health status of the insured but also on its probable evolution. Thus, there is some evidence that chronic or recurrent diseases are predictors of poor health status (high future expenditures for sickness insurance funds), while hospitalisation due to accidents will only have a short-term impact on individual health status.

To this end, we have developed a patient classification system consisting of a small number of categories, which we called "medical adjustment categories", based on the main basic pathologies (Woolf, 1998). These adjustment categories, which we briefly describe below, are based on a clinical approach:

a) *Chronic diseases* classified the following seven groups of similar patients: ischemia (ISC), mental illnesses (MEN), drug dependence (DEB), chronic dysfunction (FOA), inflammatory diseases (IMA), degenerative disease (DEA and DED);
b) *Recurrent diseases*: some infections (INF);
c) *Progressive diseases*: malignant neoplasms (TUM);
d) *Acute health problems*: accidents (ACC), acute dysfunctions (FOB), maternity (MAT).

The main limitation of such a patient classification system is that the groups

are independent of the involved organs and of the severity of the illness. A simple way to capture such additional information is to use three additional classes, identifying very high (CC2), high (CC1) and low (LOW) risk morbidities. Two surgical groups have been added to distinguish patients with major (MAJ) or minor (MIN) operations. The patients that were not assigned to any of the above categories are classified under a general category denoted OTH.

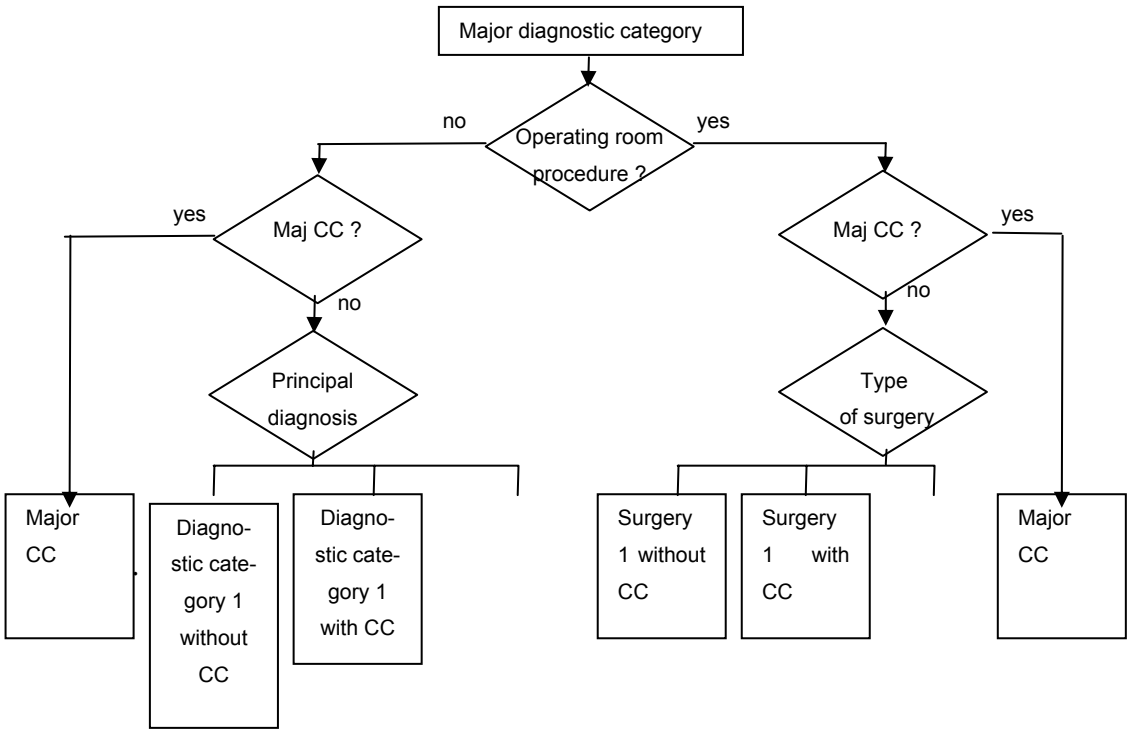
5.2 PATIENT CLASSIFICATION SYSTEMS BASED ON AP-DRGS AND SQLAPE

When this study was conducted, the main patient classification system in use in several cantons was the All-Patients Diagnosis Related Groups (AP-DRGs, version 12). However, as we explain below, we were aware of the possible limitations of AP-DRGs for the purpose of developing a risk adjustment model. For this reason, we also decided to use another PCS, known as SQLape, which is still in the development phase, merely as an illustration of an alternative to AP-DRGs.

The AP-DRGs is a categorical clinical model in which each inpatient encounter is assigned to a single, mutually exclusive risk group, using main diagnoses and operating room procedures. Developed in the late 1970s, the family of Diagnosis-Related Groups (DRGs) are now implemented worldwide, often with national adaptations (Averill et al., 1999; Jackson et al. 2001; Kroneman et al., 2001; Evers et al, 2002). The AP-DRGs logic is based on a multiphase process (or algorithm) (see Figure I). The first phase of the classification is based on the main diagnosis, specifying one of 24 major diagnostic categories corresponding more or less to medical specialities. The second phase in the algorithm is related to the presence or the absence of operating room procedures. More sophisticated classification rules were introduced in the subsequent phases to construct clinical homogeneous groups from the perspective of a diagnosis or operating room procedures. Other phases in the algorithm introduced subclasses, splitting some groups when co-morbidities or complications (CC or major CC) are identified. This is currently the only patient classification system which is compatible with routinely available hospital records in Switzerland (ICD-10 and CHOP nomenclatures). Even if it has been mostly used for cost accounting and hospital financing, it

provides an interesting way to classify patients for the purpose of developing a risk adjustment model.

FIGURE I: AP-DRGS SIMPLIFIED CLASSIFICATION SCHEME



CC = co-morbidity or complication

The mapping of the AP-DRGs in the medical adjustment categories is described in Table I below and with more details in Appendix A. A large number of AP-DRGs can easily be attributed to the main medical categories. For instance, an acute myocardial infarct is obviously classified in the ischemia category, while neuroses are clearly mental diseases. However, for other AP-DRGs, mainly groups with operating procedures, the classification is more difficult because the underlying illness is not known. If the risk adjustment model we suggest were to be retained, it is advisable to test alternative mappings of the AP-DRGs. in the medical adjustment categories. In all cases, the presence of CC is useful to identify very high (major CC=CC2) or high (other CC=CC1) risk morbidities. Similarly, high and low risk procedures are easily identified by AP-DRGs with major operating room procedures (MAJ) or other operating room

procedures (MIN)[19].

The use of AP-DRGs has another limitation as far as hospital medical data reduction is concerned, because most secondary illnesses are not considered as classifiable information. In fact, these co-morbidities (not defined as the main diagnosis) provide essential information, since they are often linked to chronic health conditions which are most predictive of future health expenditures. From this point of view, it is advisable to test other patient classification systems, based on the same routinely available inpatient encounter data but taking more explicit account of all secondary non-acute diagnoses. Merely as an illustration of an alternative to AP-DRG we present results based on a PCS, known as SQLape, which is still in the development phase. However, we thought it interesting to use SQLape as an alternative mapping in the medical categories described above.

The original aim of the SQLape classification system was to adjust for health quality measurement indicators (avoidable re-admissions, premature death, inappropriate hospital days, intensive care use), but preliminary results also show high performances for predicting hospital expenditures, using exactly the same medical information as AP-DRGs. Based on about 200 diagnostic groups and 150 procedures groups, SQLape provides for each hospital stay the medical categories listed above, as well as the etiology of the illnesses (acute, chronic, delivery, malignant, recurrent, complication, etc.) and the involved organs. One obvious advantage of this classification system is that it takes into account all pathologies (secondary diagnoses are also classified in homogeneous clinical

---

[19] Antioch and Walsh (2000) (2002) developed a disease specific (Cystic Fibrosis) Risk Adjusted Capitation Funding Model (RACFM) using Australian DRGs in the context of Victorian casemix funding. They used data for DRG 173 and estimated a linear regression model based on the variables: age, sex, emergency admission, outlier on Length of Stay, Complexity measure of the DRG grouper, number of diagnoses and number of procedures. Of these variables, emergency, outlier, complexity and procedures were significant. The aim of their study was different to ours. However, Antioch and Walsh (2003) advocate the use of the Diagnostic Cost Group/ Hierarchical Condition Category classification system for a situation similar to ours. The Victorian government, through its Committee the Risk Adjustment Working Group (RAWG) chaired by Antioch, is considering Risk Adjusted Capitation Funding Models (RACFM) for State-wide referral services for chronic diseases and also for extended episode of care arrangements (Antioch and Walsh, 2003). They advocate DCG/HCC to risk adjust national reinsurance pool arrangements and for indexing the National Australian Health Care Agreements (AHCA) that finance hospitals in all States and Territories.

groups). A second advantage is due to the fact that these pathologies are documented even if the patient has undergone an operation. The mapping from SQLape in the medical adjustment categories is described in Appendix B.

Once again, it is important to bear in mind that the only purpose of using the SQLape grouper in our study is to illustrate the potential of improvement that could be expected from using alternative classification systems. Of course, other patient classification systems, such as for example PIPDCG or, DCG if outpatient encounter data were made available, are certainly worth testing in the future.

TABLE I: MEDICAL ADJUSTMENT CATEGORIES: BRIEF CONTENT DESCRIPTION

Category	Examples
Accident (ACC)	mainly musculo-skeletal trauma
Degenerative diseases (DEA)	multiple sclerosis
	cerebro-vascular diseases
	interstitial lung diseases
	atherosclerosis
	medical back problems
	skin ulcers
	diabetes
	etc.
Drug dependency (DEB)	opioïd, cocaïn, alcohol abuse or dependency
Renal insufficiency (DED)	terminal insufficiency, mainly renal
Dysfunction (FOA)	Seizure and headache
	dysequilibrium
	circulatory disorders
	cardiac arrhythmia
	endocrine disorders
	allergic reactions
Coagulation disorders (FOB)	coagulation disorders
Inflammatory disorders (IMA)	bronchitis and asthma,
	inflammatory bowel, oeso/gastro/enteritis,
	bursitis, myositis, tendonitis
Infection (INF)	viral meningitis
	epiglottis, otitis, laryngothracheitis,
	endocarditis
	osteomyelitis
	cellulitis
Ischemia (ISC)	precerebral occlusions
	acute myocardial infarcts
	angina pectoris
	chest pain
Low risk morbidity (LOW)	DRGs: see Appendix A
	SQLape : Uterus, skin, ORL non acute conditions
Major procedures (MAJ)	major procedures (DRGs)
	brain, bowel, stomach, lung major procedures (SQLape)
Maternity (MAT)	obstetrical conditions
Minor risk procedures (MIN)	DRGs: see Appendix A
	SQLape: minor uterus, skin, ORL procedures,
	appendectomy
Mental disorders (MEN)	depression
	neurosis
	psychoses
Neoplasms (TUM)	Nervous neoplasm
	ORL malignancy
	respiratory
	digestive malignancy, etc.
Co-morbidities or complications (CC1)	DRGs : any group with "co-morbidity or complication (CC)" mention
	SQLape: mainly other digestive, peripheral nerves, respiratory, blood non acute conditions
Major co-morbidities or comp. (CC2)	DRGs : any group with "major co-morbidity or complication (CC)" mention
	SQLape: mainly other cerebral, renal, peripheral vessels, spinal cord non acute conditions

6 DATA AND METHOD

6.1 STUDIED POPULATION

All people with a health insurance policy with CSS, Helsana/Progres and KPT/CPT sickness insurance funds were included in the source population, if they were living in the cantons of Geneva, Vaud or Zurich during the 1997 to 2002 period. However, not all insurance policies were retained. The eligible insurance policies were those which were likely to give rise to future expenditures. This subset cannot thus include deceased policyholders or those who have moved to another canton or abroad. In addition, people younger than 18 were excluded because they are not intended to be included in the risk adjustment model presently in force.

In this study, we used the term “index contract” to designate the eligible contracts. The prediction of future sickness insurance funds expenditures suppose that the index contract can be followed up at least during the next year. All potentially predictive variables, including medical characteristics of inpatient encounters, should be available for the “index year”, i.e. the year of the index contract.

Since KPT/CPT data were not complete (dates of birth and years 2001/2002 missing), they were not used in this study. Similarly, the study did not involve the population living in the canton of Geneva, because we did not obtain the authorisation to use medical data from the University Hospital. Insurance data from 1997 were not used, because matching variables (particularly the date of birth) were not available in most participating hospitals. Medical data from the University Hospital Zurich were available for the years 2001 and 2002; they have been used for validation.

Finally, the studied population was restricted to:

a) People living in the canton of Vaud with a health insurance policy with CSS, Helsana/Progres during the time interval 1998 to 2001;

b) People living in the canton of Zurich with a health insurance policy with the same sickness insurance funds in 2001.

In consequence, policyholders were eligible if they satisfied the following criteria:

- at least 17 years old (year of insurance minus year of birth > 16);
- no death during the index year;
- had a policy for at least 11 months during the index year (no bias for possible hospital stays);
- had a policy for at least 11 months (same company) during the next year or death occurring during the next year (no bias for sickness insurance funds expenditures).

1,056,103 individuals satisfied the first three criteria, of whom 846,482 also satisfied the fourth criteria (363,353 for the canton of Vaud, 483,129 for the canton of Zurich).

6.2 DATA DESCRIPTION

Three records were merged in a unique relational database:

- annual insurance policies, with the following variables for each insured person: anonymous identifier, date of birth, gender, calendar year, co-payments amounts in the form of deductible and co-insurance, ambulatory expenditures, hospitalisation expenditures;
- inpatient encounters by periods recorded by sickness insurance funds: anonymous identifier (insured); hospital fee number, start date of the period, end date of the period;
- Medical data encounters recorded by hospitals: date of birth, gender, diagnoses, procedures, admission date, discharge date, discharge

status, newborn weights.

Inpatient encounters are related to hospital stays. However, we considered an additional demographic factor, namely the institutional status “nursing home” (N). The purpose was to test whether it was necessary to introduce an explicit, separate risk adjustment payment for an insured person who is resident in a registered nursing home, or if the risk adjustment model, excluding the N status, predicts mean expenditures accurately for the insured population resident in a nursing home as a whole.

Insurance and hospital data were matched using the following data: birth date, gender, hospital identifier, and entry and discharge date without considering the other variables. After matching, all potentially identifier variables (day of birth, month of birth, admission and discharge dates) were destroyed in accordance with data protection rules. If many data extractions were made, the most recent were retained.

6.3 MATCHING RULES

Matching insurance and hospital data is based on a semi-automated algorithm, which takes into account some issues relative to data quality. First, the day and month of the birth were missing in certain hospitals for some years in the canton of Vaud. The cantonal office of statistics (SCRIS) set these missing dates at 30th June of the corresponding birth year. There is an exception is for the hospital in Payerne (only for 1998 data), which set these birth dates at 1st January. Second, only the month and the year were documented for people insured with Helsana/Progres. Third, the day and month of birth were sometimes mutually exchanged.

The list of hospitals is not identical for medical records (list of hospital sites) and for insurance (fees references). A common, aggregated list of hospitals (juridical entities) has been compiled in order to facilitate record matching. Gender and date formats were also unified.

Another issue relates to the definition of a hospital stay. On the one hand, insurance records generally cut hospitalisation periods at the end of each

month. On the other hand, hospital records often cut successive stays, especially if the hospital (juridical entity) includes a rehabilitation centre. Consequently, the successive insurance periods and hospital stays were aggregated to build similar hospitalisation. The time interval between successive periods had to be less than two days.

From a theoretical point of view, all matching variables (year of birth, month of birth, day of birth, gender, hospital organisation, admission date, discharge date) should be identical. From a practical point of view, it was necessary to follow more detailed rules (see below). Only univocal matching records were retained if, for instance, a hospital record corresponds to more than one insured person. Records were matched if all matching variables, except one, were identical. After careful manual verification, the following exceptions were made:

- Exchange of day and month of birth;
- The necessity of having the same hospital was removed, if the fee was produced by the fee centre (CEESV);
- The necessity of having the same day of birth if the insurance company was Helsana/Progres (only the month and the year were given);
- The necessity of having the same day and month of birth if the hospital was the Hôpital intercantonal de la Broye and if the day of birth was set at 1st January.
- The necessity of having the same day and month of birth if the hospital was the Hôpital intercantonal de la Broye and if the day of birth was set at 30th June;
- Day of birth and discharge date not identical (often corresponding to multiple hospitalisation in the same hospital organisation);
- Day of birth and admission date not identical (idem).

Second, the insurance costs for the subsequent year were extrapolated from the previous insurance year data using insured people as the identifier.

Among the studied population, 11.4% (42,153/368,956) were hospitalised at least once during the year. 93.2% of this group have at least one corresponding medical record (39,296/42,153).

6.4 DEPENDENT VARIABLE

The dependent variable represents expenditures which are sickness insurance funds' payments to providers for both inpatient and outpatient services. Co-payments in the form of deductibles and co-insurance for the covered services that are the responsibility of the insured persons were excluded from expenditures. Because the aim of the model is to adjust for financial risks taken on by sickness insurance funds, it is logical to exclude out-of-pocket payments in the form of deductibles and co-insurance.

6.5 INDEPENDENT VARIABLES

The independent variables are following:

AGE, computed as the year of insurance contract - the year of birth

GENDER (dichotomous variable, 0= female, 1= male)

AMBULATORY COSTS (AC) (ambulatory expenditures of sickness insurance funds for the previous year)

HOSPITAL COSTS (HC) (hospital expenditures of sickness insurance funds for the previous year)

HOSPITAL STAY (H) (dichotomous variable H, yes = 1, no = 0)

NURSING HOME STAY (E) (dichotomous variable E, yes = 1, no = 0)

MEDICAL ADJUSTMENT CATEGORIES: 18 classes defined above (Section 5).

MEDICAL INFORMATION AVAILABLE (NI): dichotomous variable NI, 1 if *NO* medical information available, 0 otherwise. In some instances for ease of interpretation we will refer to the variable MED: 1 if medical information available, 0 otherwise

Prior ambulatory and hospital expenditures of sickness insurance funds are considered separately, as they can use them to predict future total expenditures. The list of hospitals and nursing homes in the cantons of Vaud and Zurich is given in Appendix C.

6.6 "ONE-YEAR" AND "MULTI-YEAR" MODELS

In our study we consider two types of adjustment models, which we called "one-year model" and "multi-year model".

The one-year model consists in predicting expenditures (i.e. sickness insurance funds' payments to providers) with all available data from the previous year (gender, age, hospital and nursing home stays, and medical adjustment categories). The medical adjustment categories are obtained alternatively as the image of the mappings from AP-DRGs and SQLape in the patient classification system, which contains medical adjustment groups.

The "multi-year" model consists in predicting the expenditures of sickness insurance funds with available data from prior information from all previous years (maximum of four years in our study). Most chronic illnesses can be cumulated over many years but some conditions do not have a "multi-year" effect, such as trauma or low risk procedures. The detailed description of multi-year variables is described in Table II.

TABLE II: AGGREGATION RULES

Type of aggregation	Rule of aggregation	Involved groups
Short term impact	Var1: 1 if condition present in current year	ACC, IMA, INF, MIN, LOW
Short term impact + cumulative effect	Var1: 1 if condition present in current year Var2: sum of occurrences in previous years	H, E, CC1, CC2, DED, FOA, MEN, MAJ, TUM, OTH
Short term impact + long-term effect	Var1: 1 if condition present in current year Var2: 1 if condition present in any of previous years	DEA, DEB, FOB,
Flexible rule	Indicators for presence of condition in each of the previous years	MAT, ISC, NI

Four different rules were used. As already mentioned some disease groups have only a short term impact which is captured by the corresponding indicator. Conditions that have long term effects are classified into two groups and the effects are captured by two variables: one group in which the effect is cumulative and is captured by the indicator of the condition in the current year and the

number of times the condition occurred in previous years; and a second group in which there is only a distinction between a recent condition and a lasting or recurring condition and in which the effect is captured by the indicator of the condition in the current year and the indicator of whether the condition occurred in *any* of the previous years. Finally the last group is one of heterogeneous behaviour through time

6.7 Descriptive Statistics

The descriptive statistics for the main variables in our study, for the canton of Vaud, are contained in Table III, IV and V. We note that our sample excludes insurees that deceased during the period studied. This biases our health statistics towards a healthier population and will have an impact on the interpretation of our results. We will return to this point in the following sections.

In Vaud, for the insurance companies considered, 11% of the insured population is hospitalized every year. Of these 11% we were unable to obtain medical statistics for about 15% (24% in 1998).

We also point out that some of our medical categories have very few observations (DED, FOB and INF for the SQLape classifier), so our results for these categories could be sensitive to the introduction of information concerning other insurance firms or other cantons and should be interpreted with care.

TABLE III: DESCRIPTIVE STATISTICS- ALL SAMPLE

Var/Year	1998	1999	2000	2001
	Sample - All			
N	95189	97047	88610	82507
tot cost at t+1	3913	4030	4280	4284
amb cost at t	2286	2384	2613	2721
Hosp cost at t	960	1032	1073	1059
Age	28.7	28.5	29.9	30.5
Male	0.410	0.411	0.405	0.404
e at t	0.010	0.012	0.012	0.014
h at t	0.110	0.110	0.111	0.111
	Sample - H=1			
N	10498	10649	9802	9131
tot cost at t+1	10464	10845	11299	11147
amb cost at t	5605	5920	6321	6657
Hosp cost at t	7341	7636	7743	7735
Age	35.9	35.8	37.4	38.1
Male	0.345	0.342	0.333	0.336
e at t	0.033	0.045	0.047	0.052
ni at t	0.241	0.143	0.162	0.134

TABLE IV: DESCRIPTIVE STATISTICS- HOSPITALIZED WITH MEDICAL INFORMATION – AP-DRG MAPPING

Var/Year	1998	1999	2000	2001
DRG				
Sample - H=1 and NI=0				
N	7964	9125	8214	7909
tot cost at t+1	10414	10615	10462	10770
amb cost at t	5569	5927	6203	6606
Hosp cost at t	7695	7551	7390	7527
Age	35.8	35.6	36.4	37.3
Male	0.348	0.345	0.326	0.335
e at t	0.029	0.041	0.033	0.042
cc1 at t	0.121	0.130	0.134	0.132
cc2 at t	0.053	0.048	0.052	0.064
acc at t	0.028	0.026	0.026	0.030
dea at t	0.179	0.179	0.168	0.174
deb at t	0.021	0.021	0.020	0.019
ded at t	0.001	0.001	0.001	0.001
foa at t	0.110	0.112	0.106	0.111
fob at t	0.007	0.008	0.008	0.008
mat at t	0.096	0.118	0.107	0.096
Men at t	0.059	0.054	0.069	0.070
ima at t	0.051	0.053	0.051	0.051
inf at t	0.029	0.029	0.033	0.028
Isc at t	0.045	0.050	0.047	0.045
maj at t	0.067	0.072	0.087	0.086
min at t	0.297	0.293	0.305	0.307
Tum at t	0.036	0.044	0.041	0.041
Low at t	0.069	0.075	0.077	0.075
Oth at t	0.154	0.126	0.115	0.118

TABLE V: DESCRIPTIVE STATISTICS- HOSPITALIZED WITH MEDICAL INFORMATION – SQLAPE MAPPING

Var/Year	1998	1999	2000	2001
SQLape				
	Sample - H=1 and NI=0			
N	8565	9486	8836	8275
tot cost at t+1	10852	10791	11242	11159
amb cost at t	5707	6002	6400	6712
hosp cost at t	7828	7691	7795	7865
Age	36.4	36.0	37.5	38.1
Male	0.351	0.345	0.335	0.341
E at t	0.036	0.046	0.047	0.052
cc1 at t	0.201	0.200	0.241	0.250
cc2 at t	0.077	0.070	0.087	0.088
Acc at t	0.041	0.040	0.040	0.045
Dea at t	0.023	0.023	0.021	0.022
Deb at t	0.070	0.078	0.078	0.084
Ded at t	0.004	0.004	0.004	0.005
foa at t	0.106	0.098	0.102	0.107
fob at t	0.005	0.004	0.009	0.017
mat at t	0.089	0.114	0.098	0.091
men at t	0.154	0.149	0.178	0.182
Ima at t	0.017	0.016	0.017	0.017
inf at t	0.009	0.008	0.009	0.009
Isc at t	0.022	0.021	0.020	0.022
Maj at t	0.029	0.028	0.033	0.039
Min at t	0.026	0.024	0.027	0.027
Tum at t	0.074	0.076	0.077	0.083
Low at t	0.036	0.035	0.028	0.034
Oth at t	0.316	0.308	0.291	0.269

7 DEVELOPMENT OF A PROSPECTIVE HEALTH-BASED RISK ADJUSTMENT MODEL: A PRELIMINARY STEP

As already mentioned, the main aim of our study is to develop a "one-year" and a "multi-year" prospective risk adjustment model. We now formally describe both models. The one-year model consists of predicting sickness insurance funds payments to providers, in the current year (t) for the next year (t+1) with data hospitalisation available in year (t). If we denote by y_{nt+1} the observed payment made on behalf of patient n at time $t+1$, x_{nt} the vector of patient n characteristics at time t that change with time (age, health information and health care utilization at time t), by z_n the vector of patient n characteristics that don't change with time (in our case the gender variable) and by u_{nt}^o a disturbance term, then we can write our one year model as:

$$y_{nt+1} = x'_{nt}\,\beta^o + z'_n\,\alpha^o + u_{nt}^o$$

If we further denote by Y_{t+1} the vector of all observations of y_{nt+1}, for all n, and use a similar notation for x_{nt}, z_n and u_{nt}^o, our model in matrix notation can be written as:

$$Y_{t+1} = X_t\beta^o + Z\alpha^o + U_t^o$$

We have data from year 1998 (t=1) up to year 2002 (t=T=5), therefore we could in principle estimate separately four of the above systems. The model we estimate is obtained by stacking the data of 4 consecutive years and is formally written as:

$$\begin{bmatrix} Y_2 \\ \vdots \\ Y_T \end{bmatrix} = \begin{bmatrix} X_1 & Z \\ \vdots & \vdots \\ X_{T-1} & Z \end{bmatrix} \begin{bmatrix} \beta^o \\ \alpha^o \end{bmatrix} + \begin{bmatrix} U_1^o \\ \vdots \\ U_{T-1}^o \end{bmatrix}$$

We also test this stacked model against the alternative model which has different coefficients for each year:

$$\begin{bmatrix} Y_2 \\ Y_3 \\ \vdots \\ Y_T \end{bmatrix} = \begin{bmatrix} X_1 & Z & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & X_2 & Z & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & X_{T-1} & Z \end{bmatrix} \begin{bmatrix} \beta_1^o \\ \alpha_1^o \\ \beta_2^o \\ \alpha_2^o \\ \\ \beta_{T-1}^o \\ \alpha_{T-1}^o \end{bmatrix} + \begin{bmatrix} U_1^o \\ U_2^o \\ \vdots \\ U_{T-1}^o \end{bmatrix}$$

The Chow test does not reject the stability of coefficients over time so whenever possible we use the stacked model. Detailed results will be presented below.

In order to predict payments to providers in year (T+1), i.e. year 2003, we use the estimates of the coefficients β^o and α^o obtained using the stacked model and data for the years 1998-2002 and set our prediction to be:

$$\hat{Y}_{T+1} = \begin{bmatrix} X_T & Z \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\alpha} \end{bmatrix}$$

The multi-year model is in all similar to the one year model but we redefine the variables x_{nt} to include information available in year (t) and in previous years, according to the description made in chapter 6 in Table II. The two year model can then be stated formally as:

$$y_{nt+1} = \tilde{x}'_{nt}\,\beta^o + z'_n\,\alpha^o + u_{nt}^o$$
$$\tilde{x}_{nt} = f(x_{nt}, x_{nt-1})$$

or in matrix notation:

$$\begin{bmatrix} Y_3 \\ \vdots \\ Y_T \end{bmatrix} = \begin{bmatrix} \tilde{X}_2 & Z \\ \vdots & \vdots \\ \tilde{X}_{T-1} & Z \end{bmatrix} \begin{bmatrix} \beta^o \\ \alpha^o \end{bmatrix} + \begin{bmatrix} U_1^o \\ \vdots \\ U_{T-1}^o \end{bmatrix}$$

The three year and four year models can be written in the same way with $\tilde{x}_{nt} = f(x_{nt}, x_{nt-1}, x_{nt-2})$ and $\tilde{x}_{nt} = f(x_{nt}, x_{nt-1}, x_{nt-2}, x_{nt-3})$ respectively. Prediction in these models proceeds a way analogous to the one year case.

Since data for 2003 is not yet available, and one would like to judge the predictive ability of different models presented we also estimate out of sample R^2 by first estimating all the models using data up to year (T-1), using the estimated model to predict payments for year (T) and then compare the prediction for year (T) with the actual value. In formal terms: we use the following model to estimate the parameters:

$$\begin{bmatrix} Y_2 \\ \vdots \\ Y_{T-1} \end{bmatrix} = \begin{bmatrix} X_1 & Z \\ \vdots & \vdots \\ X_{T-2} & Z \end{bmatrix} \begin{bmatrix} \beta^o \\ \alpha^o \end{bmatrix} + \begin{bmatrix} U_1^o \\ \vdots \\ U_{T-2}^o \end{bmatrix}$$

and predict using:

$$\hat{Y}_T = \begin{bmatrix} X_{T-1} & Z \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{\alpha} \end{bmatrix}$$

and then compute:

$$R^2 = 1 - \frac{\sum_{n=1}^{N} (y_{nT} - \hat{y}_{nT})^2}{\sum_{n=1}^{N} (y_{nT})^2}$$

We shall develop and analyse fully these types of prospective, health-based risk adjustment models in Section 8.

In this section, however, we present some intermediary models based on what is currently implemented and compare it to our “One Year” model. The risk adjustment currently in place as well as the proposition put forward by Beck can be seen as particular cases of the model we define as “One Year” therefore it makes sense to make an initial comparison between these models. The comparison will be in terms of their level of accuracy, or predictive power as

reflected by the out-of-sample R^2

7.1 INTERMEDIARY RESULTS FOR THE ONE-YEAR MODELS

We shall first present in detail the results for the canton of Vaud, then those for the canton of Zurich. Finally, we shall briefly compare the two.

7.1.1 Results for canton Vaud

The medical data from hospitalisation enabled us to allocate the patients to the 17 medical adjustment categories, or risk groups, defined in Section 5 (ACC, CC1, CC2, DEA, etc.). Three more variables are included: H, NI and E

7.1.1.1 Development of a health-based risk adjustment model

The present risk adjustment model is only based on age, gender and canton of residence. In addition, risk adjustment is presently calculated retrospectively rather than prospectively. The “demographic risk adjustment model” that we consider below (Table VI col. "Actual") is very close to the model which is presently used in Switzerland. The main differences are the following: we did not introduce a regional variable and we did not apply the model retrospectively.

TABLE VI: ACTUAL VERSUS ONE-YEAR MODELS

	Actual		Insurance		Beck Type		One Year - DRG		One Year - SQLape	
M 26-30	351	**	129		306	**	272	*	276	*
M 31-35	607	**	162		547	**	532	**	511	**
M 36-40	698	**	196	*	614	**	576	**	551	**
M 41-45	1,002	**	265	**	890	**	848	**	793	**
M 46-50	1,378	**	420	**	1,177	**	1,128	**	1,065	**
M 51-55	2,105	**	647	**	1,798	**	1,767	**	1,700	**
M 56-60	3,047	**	1,136	**	2,625	**	2,577	**	2,514	**
M 61-65	4,003	**	1,478	**	3,396	**	3,268	**	3,172	**
M 66-70	5,163	**	2,039	**	4,387	**	4,250	**	4,093	**
M 71-75	5,905	**	2,256	**	4,949	**	4,652	**	4,417	**
M 76-80	6,982	**	2,558	**	5,833	**	5,338	**	5,048	**
M 81-85	7,976	**	3,090	**	6,589	**	5,888	**	5,598	**
M 86-90	9,305	**	3,771	**	7,760	**	6,381	**	6,202	**
M 91+	11,198	**	4,575	**	9,652	**	7,629	**	7,485	**
F 18-25	855	**	225	**	608	**	811	**	815	**
F 26-30	1,822	**	403	**	1,013	**	1,609	**	1,627	**
F 31-35	1,857	**	383	**	1,127	**	1,643	**	1,675	**
F 36-40	1,656	**	331	**	1,232	**	1,426	**	1,448	**
F 41-45	1,643	**	354	**	1,351	**	1,398	**	1,408	**
F 46-50	2,146	**	516	**	1,831	**	1,843	**	1,830	**
F 51-55	2,639	**	610	**	2,261	**	2,250	**	2,263	**
F 56-60	3,262	**	852	**	2,809	**	2,777	**	2,754	**
F 61-65	3,639	**	934	**	3,165	**	3,102	**	3,101	**
F 66-70	4,355	**	1,255	**	3,750	**	3,635	**	3,626	**
F 71-75	5,317	**	1,746	**	4,560	**	4,330	**	4,228	**
F 76-80	6,943	**	2,538	**	5,932	**	5,355	**	5,273	**
F 81-85	8,249	**	3,087	**	6,997	**	5,924	**	5,804	**
F 86-90	10,444	**	4,011	**	8,976	**	7,033	**	6,919	**
F 91+	14,486	**	6,124	**	12,951	**	9,676	**	9,661	**
Ambul. Costs			0.954	**						
Hospit. costs			0.588	**						
H at t-1					6,550	**	-671	**	-862	**
e at t-1							11,094	**	10,398	**
cc1 at t-1							7,710	**	4,708	**
cc2 at t-1							9,037	**	9,782	**
acc at t-1							2,709	**	2,454	**
dea at t-1							7,802	**	5,453	**
deb at t-1							9,712	**	5,632	**

ded at t-1							35,539	**	72,960	**
foa at t-1							3,723	**	6,321	**
fob at t-1							12,607	**	16,685	**
mat at t-1							-104		648	**
men at t-1							5,010	**	9,276	**
ima at t-1							3,665	**	3,310	**
inf at t-1							5,716	**	4,153	**
isc at t-1							4,817	**	4,117	**
maj at t-1							2,736	**	8,284	**
min at t-1							3,562	**	-870	*
tum at t-1							10,057	**	10,433	**
zero at t-1							1,419	**	1,603	**
oth at t-1							3,315	**	2,673	**
ni at t-1							8,008	**	6,843	**
Constant	1,040	**	145	*	817	**	817	**	863	**
N Obs	363,353		363,353		363,353		363,353		363,353	
R-squared	0.06		0.31		0.10		0.14		0.18	

* significant at 5%; ** significant at 1%

As expected, predicted health expenditures strongly increase with age. However, it is also clear that the demographic variable only explains 6% of the variability of health expenditures.

Our general assumption is that sickness insurance funds are able to make good predictions about future expenditures based on the data available to them. These data are hospitalisation during the current year (H), present ambulatory expenditures (Ambul. Costs) and hospitalisation expenditures (Hospit. Costs). The “sickness insurance funds’ prediction model” that we consider in (Table VI, col “Insurance”).corresponds to this assumption

The regression results presented in clearly show that past and future expenditures are highly correlated. The coefficient of ambulatory is almost equal to one, reflecting an empirical rule of thumb stating that next year’s expenditures will be equal to present expenditures.

Hospitalisation expenditures however have a different structure, correlated to an amount of about 59% of present expenditures. This confirms our assumption according to which hospitalised individuals form a rather heterogeneous group, where costly stays are good predictors of future higher expenditures, while low cost stays are almost negligible.

Compared to the demographic risk adjustment model, the effect of age structure is significantly reduced; past expenditures are much better predictors than age alone. The introduction of the three additional variables - hospital stay, ambulatory costs and hospital costs - boosts the R^2, up to 31% of explained variance.

From the sickness insurance funds' prediction model, we see that the incentive for risk selection is very high. Insurers are quite well informed about future risks, and the risk adjusters currently used cover only a very small amount.

Introducing past expenditures in a risk adjustment system would obviously suppress this incentive to select. This would also ruin the insurance system, since the risk adjustment derived from the sickness insurance funds' prediction model implies that at the end of the year we pay an amount equal to the previous year's expenditure. Thus, *past expenditures should not be included as risk adjusters in the risk adjustment models*.

The first improvement to the demographic risk adjustment model is the introduction of the H indicator (hospital stay) only as an additional risk adjuster. We called this model "Beck-type risk adjustment model", because it is constructed in the same way as proposed by Beck (1999). The econometric estimation of this model is presented in Table VI, col "Beck")

The introduction of the variable H as a risk adjuster, as in the Beck-type risk adjustment model is a first step in our analysis. From the results presented in Table VI, this variable is clearly strongly related to expected future expenditures. Compared to the demographic risk adjustment model, the effect of age structure is significantly reduced, indicating that the H variable does indeed capture some part of the health status.

As Table VI shows, the risk adjustment payment scheme that would derive from the Beck-type risk adjustment model gives less consideration to the age structure than the demographic risk adjustment model. It gives a risk adjustment payment of CHF about 6,500 for each insured person who had at least one inpatient stay during the current year.

For the mean of hospitalisation expenditures this amount is close to the amount predicted by the sickness insurance funds' prediction model (.95 x CHF

6,100+.59 x CHF 7,600- CHF 2700 = CHF 7,500, where CHF 6,100 is the average ambulatory expenditure, where CHF 7,600 is the average hospital expenditure and CHF 2700 is the average excess contribution of the age structure variables in the Beck type model).

The next step in our analysis consists in introducing the reasons for the hospitalisation as an explanatory variable. To this end, we split the H variable and assign the patients to the 18 medical adjustment categories, or risk groups, defined in Section 5, using medical data from hospitalisation.

The results obtained from the mapping of the AP-DRGs in the medical adjustment categories, described in Appendix A, are presented in (Table VI , col "One Year - DRG"). Using AP-DRGs, we see that this health-based risk adjustment model does improve significantly on the previous regression: 14% of the expected variance is now described by the regression with demographic and medical adjustment categories resulting from the AP-DRGs mapping.

The interpretation of the regression coefficients presented in Table VI should be done with respect to three groups of individuals.

First, let us consider the “general population”, which consists of those for whom there were no sickness insurance funds’ payments to providers for inpatient services during the current year. For this group, the only risk adjusters are the age classes and gender, as in the present risk adjustment system. Here again, compared to the demographic risk adjustment model, the effect of age structure is significantly reduced. Thus, by using a health-based risk adjustment model which takes into account the health status of the policyholders, we could greatly improve the risk adjustment payment scheme for those who did not have any inpatient stays during the current year.

Second, let us consider the group of individuals for whom there were sickness insurance funds’ payments to providers for inpatient services during the current year (H=1). But for whom we could not obtain any further information (NI=1). For this sub-group, all the remaining risk group indicators are equal to zero. Not surprisingly, the value of the coefficient is close to the value obtained in Beck-type risk adjustment model; the risk adjustment payment is equal to a fixed bonus of CHF 7,337 (-671+8008=7337) in the case of no AP-DRGs mapping in the medical adjustment categories.

Finally, there are all patients with at least one documented inpatient stay. For these, we have H=1, and medical information NI=0. They were assigned to the medical adjustment category, or risk group (or groups), corresponding to their health status. We should note that for a lot of conditions all these variables are set at zero, while it may happen that some other patients figure in more than one medical adjustment category.

It is interesting to consider in more detail the results for the sub-group of individuals for whom we have H=1 and NI=0. For individuals of the same age and gender, the additional risk-adjusted payments to the sickness insurance funds are equal to sum of the coefficients of the variables H and the risk group corresponding to their health status. For example, for a normal delivery which is coded as MAT=1, the risk adjustment payment amount is CHF -775 (-671-104= -775).

This value is relatively small and negative; an inpatient stay for delivery is not a predictor for future expenditures. On the contrary, the negative sign expresses the fact that, on average, delivery in the current year is a predictor of good health in the next year.

On the other hand, the additional risk adjustment payment for a patient with terminal insufficiency (DED) is equal to -CHF 671+CHF 35,539=CHF 34,868. We nevertheless remind the reader that this category in present in very few cases.

The MIN (minor intervention) class was introduced in order to take into account interventions which help to cure the patient. A patient in this group ends up with a risk adjustment payment amount of CHF 2,891

The two groups of severity CC1 and CC2 are added to the risk adjustment and the payments amount to CHF 7,039 and CHF 8,366 respectively.

Clearly, the results we have just presented show that, using medical information from prior hospitalisation, as we did in the case of the health-based risk adjustment model constructed from AP-DRGs, makes a great difference, not only compared to the demographic risk adjustment model, but also compared to the Beck-type risk adjustment model. This is because the amount of risk adjustment payments varies widely depending on the reasons for the hospitalisation. In particular, in this health-based risk adjustment model, deliveries should not automatically mean a risk adjustment payment, while the

risk-adjusted payment for policyholders within the DED medical category should be about CHF 35,000.[20]

It is important to observe that the nursing home variable is highly significant. We thus found it necessary to include an explicit, separate risk adjustment payment for the "nursing home" variable. The purpose of this risk adjuster is to ensure that the risk adjustment model takes into account the average expenditures of the insured population resident in a registered nursing home.

To illustrate the potential improvement that could be achieved from alternative classification systems which explicitly take into account the severity of the illness, we reproduce below the results obtained on the basis of the present version of SQLape (see Section 5). The results obtained from using the mapping of SQLape in the medical adjustment categories, described in Appendix B, are presented in Table VI, col "One Year – SQLape".

Using SQLape, we see that the health-based risk adjustment model is an improvement on the previous regression: 18% of the expected variance is described by the regression with demographic and the medical adjustment categories resulting from the SQLape mapping, compared to 14% for the AP-DRGs mapping.

Let us now compare the risk adjustment payment implied by the SQLape mapping, compared to that implied by the AP-DRGs mapping. For example, with the SQLape mapping, for a normal delivery (MAT), the risk adjustment payment is CHF -214.

This value is smaller than in the case of AP-DRGs mapping. Here, this value is not significantly different from zero.

In contrast, the additional risk adjustment payment for a patient with terminal insufficiency (DED) is equal to CHF 72,098, which is almost twice the amount of the AP-DRG-based risk adjustment payment.

A patient in the MIN (minor intervention) medical category group ends up with a risk adjustment payment of –CHF 1,732. Note the minus sign, indicating that the construction of this class was successful; if we apply the adjusters as they are, the minor interventions are indicators of lower expenditures for the next

[20] Note that, as a result of the DCG algorithm presented earlier, deliveries are also assigned in the

year.

The two groups of severity CC1 and CC2 are added to the risk adjustment and amount to 3,846 and 8,920 respectively. The risk adjustment payment for the CC1 group of severity is almost half the total payment resulting from the AP-DRGs.

It should also be noted that the nursing home variable is highly significant, with a risk adjustment payment total similar to the previous model.

As expected, the results presented in Table VI reinforce the importance of setting up an appropriate health-based risk adjustment model using medical information from prior hospitalisation.

They provide an additional incentive to test other patient classification systems which explicitly take into account the severity of illness and which are mainly diagnosis-based and exclude treatments or procedures as far as possible.

7.1.1.2 A polynomial version of the age risk groups' structure

During the development of the groupers and the risk adjustment model, it became obvious that the current structure of age risk groups is not optimal, particularly the final open class of 91+. A smaller number of age risk groups would lower the precision of the computation of the average per class. Therefore, we developed a polynomial version of the age risk groups' structure. We ended up with a polynomial in the order of 5 as a better trade-off between accuracy and over-fitting. The evolution of the risk adjuster which corresponds to age is approximately the same as the one derived from the present risk groups' structure, but without the gaps between age risk groups, and with a continually rising risk adjustment payment for policyholders aged 91 and more.

We estimated a health-based risk adjustment model similar to the one represented in Table VI, except that now the age risk groups have been replaced by their polynomial versions. The results of this estimation are presented in

---

lowest PIDCGs corresponding to zero expected future costs [see Pope et al. (2000), Table 4].

Table VII.

TABLE VII: POLYNOMIAL VERSIONS OF THE AGE RISK GROUPS

	One Year - DRG		One Year - DRG	
M 26-30	272	*		
M 31-35	532	**		
M 36-40	576	**		
M 41-45	848	**		
M 46-50	1,128	**		
M 51-55	1,767	**		
M 56-60	2,577	**		
M 61-65	3,268	**		
M 66-70	4,250	**		
M 71-75	4,652	**		
M 76-80	5,338	**		
M 81-85	5,888	**		
M 86-90	6,381	**		
M 91+	7,629	**		
F 18-25	811	**		
F 26-30	1,609	**		
F 31-35	1,643	**		
F 36-40	1,426	**		
F 41-45	1,398	**		
F 46-50	1,843	**		
F 51-55	2,250	**		
F 56-60	2,777	**		
F 61-65	3,102	**		
F 66-70	3,635	**		
F 71-75	4,330	**		
F 76-80	5,355	**		
F 81-85	5,924	**		
F 86-90	7,033	**		
F 91+	9,676	**		
Gender			-877	**
mage_1			56	**
mage_12			-62	**
mage_13			35	**
mage_14			-5.947	**
mage_15			0.327	**
fage_1			131	**
fage_12			-118	**
fage_13			47	**

fage_14			-7.115	**
fage_15			0.399	**
e at t-1	11,094	**	11,084	**
h at t-1	-671	**	-676	**
cc1 at t-1	7,710	**	7,699	**
cc2 at t-1	9,037	**	9,039	**
acc at t-1	2,709	**	2,714	**
dea at t-1	7,802	**	7,802	**
deb at t-1	9,712	**	9,722	**
ded at t-1	35,539	**	35,539	**
foa at t-1	3,723	**	3,719	**
fob at t-1	12,607	**	12,614	**
mat at t-1	-104		-74	
men at t-1	5,010	**	5,014	**
ima at t-1	3,665	**	3,673	**
inf at t-1	5,716	**	5,723	**
isc at t-1	4,817	**	4,810	**
maj at t-1	2,736	**	2,740	**
min at t-1	3,562	**	3,563	**
tum at t-1	10,057	**	10,054	**
zero at t-1	1,419	**	1,418	**
oth at t-1	3,315	**	3,314	**
ni at t-1	8,008	**	8,019	**
Constant	817	**	1,815	**
N Obs	363,353		363,353	
R-squared	0.14		0.14	

* significant at 5%; ** significant at 1%

For ease of comparison the first column of results of Table VII reproduces the corresponding column of Table VI. The results presented in both columns of Table VII are very close, and no coefficients differ significantly. Therefore, for the multi-year regressions we shall only present the results with the polynomial age risk adjuster.

The adjustments according to age are better viewed in a graph (see Figure II and **Error! Reference source not found.**).

FIGURE II: POLYNOMIAL AGE STRUCTURE VERSUS AGE GROUPS - MALES

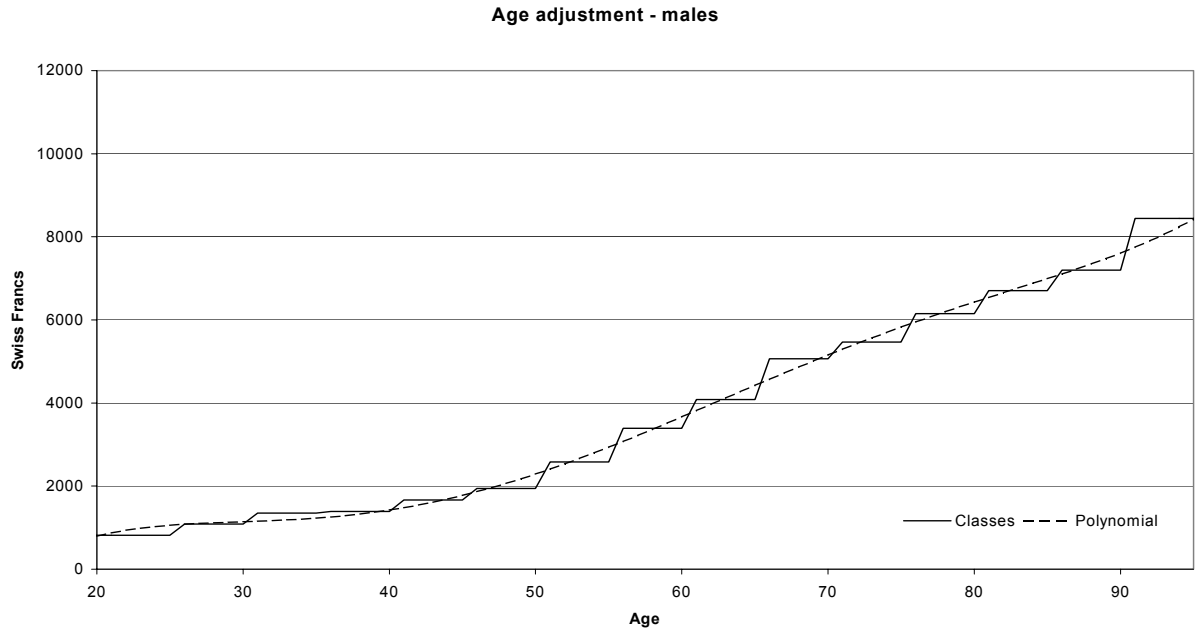
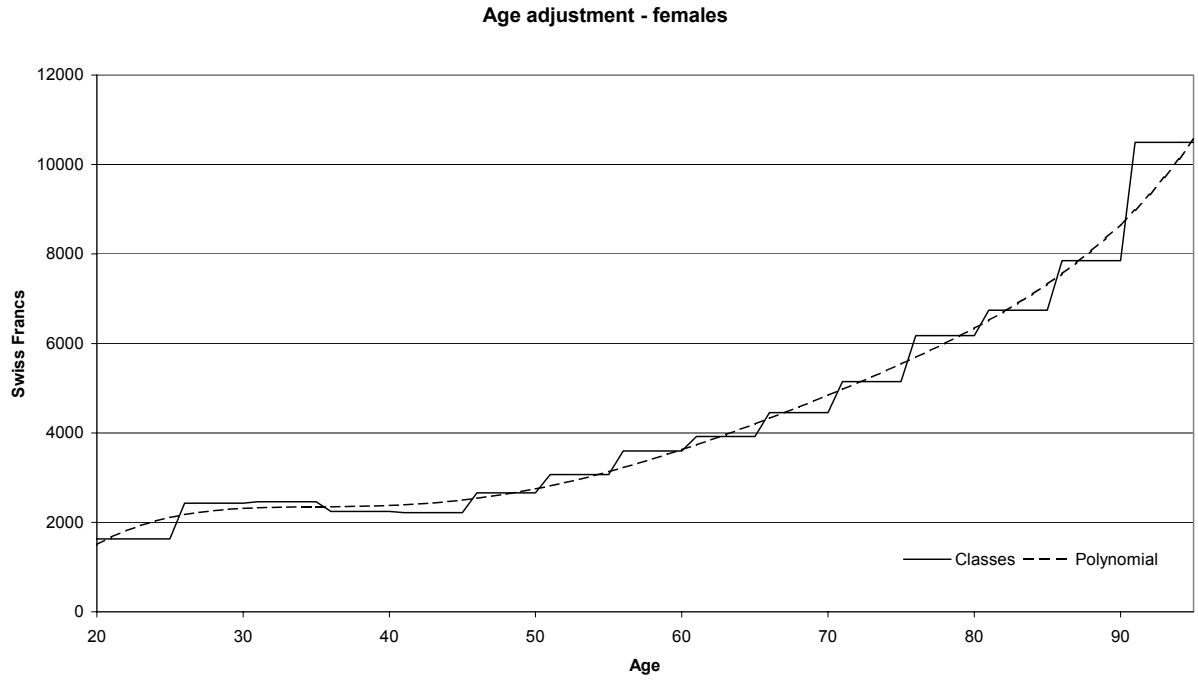


FIGURE III: POLYNOMIAL AGE STRUCTURE VERSUS AGE GROUPS - FEMALES



7.1.2 Results for the canton of Zurich

The one-year regression has been repeated with the polynomial version for policyholders in the canton of Zurich. Overall, the results are similar, even if we observe some variations in the inflammation (IMA), infection (INF) and ischemia (ISC) groups. The predictive performance is slightly lower (3% lower) for the canton of Zurich. This result is probably due to the fact that the model has been applied to all insurers in the canton of Zurich, even if medical information was available only for patients staying in the University Hospital. This probably also explains why the coefficient of major co-morbidities (CC2) is higher in this canton (more severe illnesses) and the great difference for the NI variable (availability of medical data if the patient is hospitalised).

The R2 is slightly lower for the canton of Zurich. This cross-validation is an argument in favour of the universality of the proposed model. Of course, the adjustment coefficients should be completed with data from other cantons, if such an adjustment model is to be implemented.

TABLE VIII: COMPARISON BETWEEN CANTONS VAUD AND ZURICH

	One year - DRG VD		One year - DRG ZH		One year - SQL VD		One year - SQL ZH	
Gender	-877	**	-363	**	-885	**	-369	**
mage_1	56	**	56	**	57	**	56	**
mage_12	-62	**	-50	**	-68	**	-49	**
mage_13	35	**	21	**	39	**	21	**
mage_14	-5.947	**	-2.959	**	-6.755	**	-2.831	**
mage_15	0.327	**	0.152	**	0.386	**	0.143	**
fage_1	131	**	34	**	135	**	34	**
fage_12	-118	**	-7.563		-125	**	-6.942	
fage_13	47	**	3.931		50	**	3.524	
fage_14	-7.115	**	-0.795		-7.760	**	-0.707	
fage_15	0.399	**	0.104	**	0.442	**	0.098	**
e at t	11,084	**	5,140	**	10,387	**	5,102	**
h at t	-676	**	4,245	**	-866	**	2,891	**
cc1 at t	7,699	**	3,443	**	4,707	**	2,330	**
cc2 at t	9,039	**	7,429	**	9,779	**	5,981	**
acc at t	2,714	**	3,054	**	2,442	**	2,336	**
dea at t	7,802	**	2,567	**	5,456	**	1,342	**
deb at t	9,722	**	742		5,640	**	2,638	**
ded at t	35,539	**	33,779	**	72,926	**	51,390	**
foa at t	3,719	**	-128		6,313	**	3,612	**
fob at t	12,614	**	7,740	**	16,708	**	12,971	**
mat at t	-74		-3,869	**	676	**	-1,445	**
men at t	5,014	**	22		9,279	**	5,836	**
ima at t	3,673	**	1,372	*	3,307	**	-348	
inf at t	5,723	**	3,140	**	4,166	**	6,048	**
isc at t	4,810	**	2,377	**	4,111	**	564	
maj at t	2,740	**	2,751	**	8,288	**	6,034	**
min at t	3,563	**	-257		-862	*	-2,623	**
tum at t	10,054	**	3,022	**	10,425	**	5,828	**
zero at t	1,418	**	-796		1,617	**	-1,049	
oth at t	3,314	**	-598		2,672	**	-60	
ni at t	8,019	**	-215		6,852	**	1,079	**
Constant	1,815	**	1,252	**	1,877	**	1,257	**
N Obs	363,353		483,129		363,353		483,129	
R-squared	0.14		0.14		0.18		0.15	

* significant at 5%; ** significant at 1%

7.1.3 OLS versus non-linear models

In Section 4 we described the several alternatives to OLS estimator of risk adjustment parameters. In order to compare OLS with these alternatives we performed the regressions with three alternative estimation methods: two part model, two-step (Heckman) and simple Tobit. Here, we only focus on the differences between estimation methods. The results for the model "One year - DRG" are presented in Table VII.

TABLE IX: NON-LINEAR MODELS

	Probit		Two-part		Tobit		Heckman	
Gender	-6.57E-01	**	-701	**	-2,683	**	1,789	**
mage_1	-3.80E-02	**	96	**	-72	**	271	**
mage_12	3.10E-02	**	-86	**	38		-228	**
mage_13	-9.00E-03	**	44	**	14		80	**
mage_14	1.00E-03	**	-7.55E+00	**	-3.87E+00	*	-12	**
mage_15	0.00E+00	**	4.34E-01	**	2.46E-01	*	6.91E-01	**
fage_1	-1.60E-02	**	158	**	108	**	203	**
fage_12	5.00E-03		-137	**	-115	**	-148	**
fage_13	1.00E-03		52	**	51	**	47	**
fage_14	0.00E+00		-7.87E+00	**	-8.23E+00	**	-6.32E+00	**
fage_15	0.00E+00		4.40E-01	**	4.71E-01	**	3.33E-01	**
e at t-1	6.45E-01	**	10,969	**	11,281	**	10,608	**
h at t-1	6.68E-01	**	-918	**	100		-2,479	**
cc1 at t-1	3.46E-01	**	7,655	**	7,706	**	7,794	**
cc2 at t-1	4.04E-01	**	8,962	**	9,051	**	8,980	**
acc at t-1	4.50E-02		2,739	**	2,588	**	2,776	**
dea at t-1	3.04E-01	**	7,771	**	7,800	**	7,457	**
deb at t-1	4.17E-01	**	9,720	**	10,459	**	9,680	**
foa at t-1	6.20E-02		3,725	**	3,690	**	4,311	**
fob at t-1	2.54E-01		12,637	**	12,690	**	11,148	**
mat at t-1	-3.05E-01	**	-59		-293		473	
men at t-1	2.03E-01	*	5,069	**	5,390	**	5,346	**
ima at t-1	8.60E-02		3,689	**	3,674	**	4,201	**
inf at t-1	-2.80E-02		5,886	**	5,776	**	5,849	**
isc at t-1	2.50E-01	*	4,738	**	4,786	**	5,073	**
maj at t-1	1.89E-01	*	2,639	**	2,668	**	2,548	**
min at t-1	2.50E-02		3,573	**	3,592	**	3,479	**
tum at t-1	5.97E-01	**	9,948	**	10,058	**	9,599	**
zero at t-1	-1.87E-01	**	1,572	**	1,341	**	1,729	**
oth at t-1	-1.14E-01		3,392	**	3,235	**	3,752	**

ni at t-1	-8.00E-03		8,203	**	7,986	**	8,434	**
ded at t-1			35,095	**	36,399	**	37,768	**
mymills							-9,046	**
Constant	1.13E+00	**	2,100	**	780	**	4,271	**
N Obs	363,353		307,889		363,353		237,836	
R-squared			0.13				0.14	

* significant at 5%; ** significant at 1%

We should now compare the accuracy of the estimated future expenditures obtained by these three regressions. For this purpose we computed out of sample R2 as described in the beginning of the chapter, that is we estimate each model using data from years 1998 up to 2001 and the use the estimated models to predict year 2002 expenses. Predicted expenses are compared to actual expenses. The results of this exercise are given in Table X.

TABLE X: OUT OF SAMPLE R^2 – ONE YEAR MODELS

	Out of sample R2 DRG	Out of sample R2 SQLape
Actual	.27	.27
Insurance	.56	.56
Beck Type	.31	.31
One year model	.36	.40
One year two-part model	.36	.42
One year Tobit model	.34	.38
One year Heckman selection model	.36	.40

The OLS procedure does not clearly dominate the remaining procedures; in particular the two-part model performs better for the model which uses the SQLape classification system. However the results are close enough to justify, on the basis of its simplicity, the use of OLS in an exploratory phase of the study of risk adjusters as it is clearly the case of the present report. We will therefore, proceed our analysis presenting only the results for OLS but a note is made that future work should explore potential improvements in prediction achievable by the use on non-linear models.

8 PROSPECTIVE HEALTH-BASED RISK ADJUSTMENT MODEL

The models presented in Section 7 were intended to mimic what in currently implemented and provide a simple extension based on information of previous year health care utilization. The purpose of this section is to develop a prospective health-based risk adjustment model based on these medical adjustment groups which use information on several years and to evaluate its predictive power. This prospective health-based risk adjustment model is developed by means of the two types of model which we called “one-year model” which was already presented in the previous section and “multi-year model”.

Recall that the one-year model consists of predicting expenditures (i.e. sickness insurance funds payments to providers) for the next year (t+1) with data on prior hospitalisation (gender, age, hospital and nursing home stays, medical adjustment categories) from the year (t), the current year . In turn, the multi-year model consists of predicting expenditures for next year (t+1) with data from prior hospitalisation in the current year (t) and in previous years (t-1), (t-2), (t-3), etc. In the development of the multi-year risk adjustment model, we used the set of aggregation rules described in.

8.1 DEVELOPMENT OF THE PROSPECTIVE HEALTH-BASED RISK ADJUSTMENT MODEL

The regressions are performed following the same pattern as the one-year presented in Section 7. The multi-year models based on the medical adjustment categories that we suggest in this study are presented in Table XI and Table XII: for the AP-DRGs and SQLape mappings respectively. The first column of each of these tables replicates the results of one-year models for comparison purposes. The regressions as a whole perform well; the percentage of explained variance goes up to 21 or 25% in with the AP - DRG mapping with the introduction. This last value is about three times the total derived from the demographic risk

adjusters. This however is not a good comparative measure between the different models since each of the models is estimated with different samples as described in detail at the beginning of section 7. We will return to the comparison of the predictive power of the models further below.

TABLE XI: MULTI-YEAR MODELS: DRG

	One Year - DRG		Two Year - DRG		Three Year - DRG		Four Year - DRG	
Gender	-877	**	-816	**	-806	**	-751	**
mage_1	56	**	50	*	34		33	
mage_12	-62	**	-59	*	-32		-29	
mage_13	35	**	35	**	21		19	
mage_14	-5.947	**	-5.981	**	-3.611		-2.962	
mage_15	0.327	**	0.338	**	0.197		0.144	
fage_1	131	**	128	**	126	**	105	**
fage_12	-118	**	-121	**	-116	**	-96	**
fage_13	47	**	49	**	45	**	38	**
fage_14	-7.115	**	-7.579	**	-6.833	**	-5.618	**
fage_15	0.399	**	0.426	**	0.366	**	0.294	*
e at t	11,084	**	7,453	**	6,558	**	6,305	**
sum of e before t			5,471	**	3,506	**	2,671	**
h at t	-676	**	-1,208	**	-1,057	**	-561	*
sum of h before t			1,946	**	1,784	**	1,498	**
cc1 at t	7,699	**	7,469	**	7,201	**	5,998	**
sum of cc1 before t			3,510	**	2,857	**	2,832	**
cc2 at t	9,039	**	8,515	**	8,713	**	7,248	**
sum of cc2 before t			4,319	**	4,307	**	3,592	**
acc at t	2,714	**	2,268	**	2,921	**	2,991	**
dea at t	7,802	**	7,499	**	7,709	**	7,862	**
max of dea before t			2,791	**	2,088	**	2,189	**
deb at t	9,722	**	7,016	**	5,213	**	4,205	**
max of deb before t			3,873	**	1,983	**	396	
ded at t	35,539	**	35,758	**	44,503	**	15,809	**
sum of ded before t			20,340	**	13,127	**	14,002	**
foa at t	3,719	**	3,660	**	2,868	**	2,007	**
sum of foa before t			628	**	554	**	-375	*
fob at t	12,614	**	12,283	**	11,685	**	15,653	**
max of fob before t			9,573	**	10,529	**	6,383	**
mat at t	-74		286		-71		-377	
mat at t-1			-1,954	**	-1,951	**	-1,658	**
mat at t-2					-1,699	**	-1,421	**
mat at t-3							-1,192	**
men at t	5,014	**	4,102	**	3,436	**	1,037	*
sum of men before t			1,986	**	1,546	**	1,840	**

ima at t	3,673	**	3,394	**	2,768	**	2,074	**
inf at t	5,723	**	5,483	**	5,125	**	5,060	**
isc at t	4,810	**	4,525	**	3,445	**	3,747	**
isc at t-1			628	*	391		-131	
isc at t-2					673	*	810	
isc at t-3							-522	
maj at t	2,740	**	3,299	**	3,059	**	2,879	**
sum of maj before t			-395		-137		-117	
min at t	3,563	**	3,897	**	3,519	**	3,715	**
tum at t	10,054	**	9,403	**	9,235	**	9,699	**
sum of tum before t			2,225	**	2,118	**	2,504	**
zero at t	1,418	**	1,541	**	2,269	**	1,881	**
oth at t	3,314	**	3,595	**	3,041	**	2,539	**
sum of oth before t			-360		-310		-577	**
ni at t	8,019	**	8,649	**	8,649	**	6,918	**
ni at t-1			1,342	**	1,248	**	2,259	**
ni at t-2					187		486	
ni at t-3							27	
Constant	1,815	**	1,689	**	1,595	**	1,521	**
N Obs	363,353		255,313		157,493		73,326	
R-squared	0.14		0.16		0.21		0.25	

* significant at 5%; ** significant at 1%

We now look at the temporal evolution of the risk adjustment payments obtained by estimating the prospective risk adjustment models based on the medical adjustment categories constructed from the AP-DRGs (Table XI). The coefficients remain stable for most variables in the multi-year model, with the exception DEB, DED, FOA and MEN.

In the case of FOB the coefficient does not remain stable but this reflects only a different split between short-term and long-term-effects. For instance in the two year model a patient assigned to the category FOB in the current year the insurance firm would receive a risk compensation of CHF 12,283, if the same patient had been assigned to the same category in the past then the compensation would be CHF 12,283+CHF 9,573=CHF 21,856. For the three and four year models these values are CHF 11,685+CHF 10,529=CHF 22,214 and CHF 15,653+CHF 6,383=CHF 22,036 respectively.

The case of TUM illustrates well the effect of a health condition that is modelled as a progressive condition. In the one-year formulation the

compensation is CHF 10,054. In the multi-year model the compensation for the current year is around CHF 9,500 (depending on the multiyear model considered) and there is an additional compensation of CHF 2,300 x number of past years the condition has been present. In summary, the total is CHF 9,500 if the current year is the first year where the condition is present, CHF 11,800 if the current year is the second year the condition is present, CHF 14,100 if the current year is the third year the condition is present, etc.

The case of DEA is a clear case of a stable distinction between a recent event and a past occurrence of the condition. If the condition is present in the current year the compensation is around CHF 7,600. However if the condition was present in any past year, there is an additional compensation of around CHF 2,300. This is higher than the single compensation of CHF 7,800 obtained from the one-year model.

TABLE XII: MULTI-YEARS MODELS: SQLAPE

	One Year – SQL		Two Year – SQL		Three Year – SQL		Four Year - SQL	
Gender	-885	**	-806	**	-785	**	-747	**
mage_1	57	**	52	**	39		36	
mage_12	-68	**	-68	**	-47		-42	
mage_13	39	**	39	**	29	**	26	
mage_14	-6.755	**	-7.03	**	-5.254	**	-4.492	
mage_15	0.386	**	0.413	**	0.309	**	0.248	
fage_1	135	**	129	**	126	**	108	**
fage_12	-125	**	-122	**	-114	**	-98	**
fage_13	50	**	50	**	46	**	39	**
fage_14	-7.76	**	-7.9	**	-7.053	**	-5.913	**
fage_15	0.442	**	0.452	**	0.39	**	0.316	**
e at t	10,387	**	6,816	**	6,112	**	5,961	**
sum of e before t			5,393	**	3,361	**	2,597	**
h at t	-866	**	-725	**	162		1,018	**
sum of h before t			407	*	781	**	852	**
cc1 at t	4,707	**	4,285	**	3,889	**	3,478	**
sum of cc1 before t			1,758	**	1,122	**	1,031	**
cc2 at t	9,779	**	8,733	**	7,702	**	6,187	**
sum of cc2 before t			3,757	**	3,587	**	2,484	**
acc at t	2,442	**	2,554	**	2,125	**	2,368	**
dea at t	5,456	**	5,200	**	4,559	**	3,247	**
max of dea before t			2,898	**	2,998	**	2,318	**
deb at t	5,640	**	4,544	**	3,917	**	2,527	**
max of deb before t			2,223	**	1,003	**	629	**

ded at t	72,926	**	56,543	**	47,438	**	52,299	**
sum of ded before t			40,078	**	35,095	**	21,173	**
foa at t	6,313	**	4,942	**	3,510	**	2,614	**
sum of foa before t			3,711	**	2,396	**	1,768	**
fob at t	16,708	**	13,870	**	13,140	**	13,882	**
max of fob before t			5,375	**	4,667	**	380	
mat at t	676	**	532		-598		-1,582	**
mat at t-1			-264		-744	**	-815	*
mat at t-2					-666	*	-510	
mat at t-3							-898	*
men at t	9,279	**	7,777	**	6,214	**	5,052	**
sum of men before t			4,736	**	3,619	**	2,929	**
ima at t	3,307	**	2,930	**	1,094	*	-645	
inf at t	4,166	**	3,743	**	3,055	**	2,687	**
isc at t	4,111	**	2,535	**	1,444	**	576	
isc at t-1			1,673	**	1,762	**	659	
isc at t-2					459		1,200	
isc at t-3							-1,953	**
maj at t	8,288	**	7,545	**	6,016	**	4,344	**
sum of maj before t			4,518	**	3,712	**	2,325	**
min at t	-862	*	-521		-812		-2,048	**
tum at t	10,425	**	9,402	**	8,318	**	8,314	**
sum of tum before t			4,032	**	2,967	**	2,818	**
zero at t	1,617	**	1,454	**	667		83	
oth at t	2,672	**	2,427	**	1,448	**	577	
sum of oth before t			571	**	111		-194	
ni at t	6,852	**	6,774	**	5,381	**	4,050	**
ni at t-1			2,809	**	2,263	**	2,539	**
ni at t-2					1,327	**	1,087	**
ni at t-3							794	**
Constant	1,877	**	1,753	**	1,666	**	1,611	**
N Obs	363,353		255,313		157,493		73,326	
R-squared	0.18		0.19		0.26		0.3	

* significant at 5%; ** significant at 1%

The results for the SQLape models (Table XII) are qualitatively similar. There is a considerable difference in the amount of compensation attributed to the DED category. However at mentioned commented earlier in section 6, the results for this set of conditions are based in very few observations and a minor change in classification is likely to produce diverging results.

We also note that if we test the restrictions we imposed on the multi-year models, regarding the form in which the information from previous years is

taken into account (i.e. our aggregation rules described earlier) by means of an F-test, then the only conditions for which our rules are rejected are DED and MEN.

At this point it is useful to recall that our sample excludes individuals that died during the four years considered. This has strong implication for the interpretation of the results, since we potentially excluded a population that is considerable less healthy. They results must be interpreted as the risk adjustment amounts to be had *if* the insured does not die in the following year. According to the present formulation, compensation is prospective only for insurees that do not died. For does the do, there is no conclusion that could be drawn from our calculations except that if one implements the risk adjustment model as they are presented then a retrospective compensation for the individuals the die would have to be made. We leave the development of fully prospective models that take into account mortality to future work.

8.2 VALIDATION OF THE PROSPECTIVE HEALTH-BASED RISK ADJUSTMENT MODEL

A first validation test to perform is the adequacy of the risk adjusters is to try to access the remaining selection incentives. For this purpose regress the actual future cost minus the risk adjustment against the present expenditures (Table XIII). The perfect risk adjuster should produce zero coefficients, denoting that this information is not useful to estimate future costs after the risk adjustment has been made. On the other hand, the worse risk adjuster should produce coefficients equal to the initial values for the sickness insurance funds' prediction model, namely 0.88 and 0.53 for hospital costs and ambulatory costs respectively.

Past expenditures still explain about 14% of the remaining variance. However, the coefficients are much lower than without risk adjustment. The coefficient of current ambulatory costs is 0.776 (0.68 in SQLape) instead of 0.88, while the coefficient of current hospital costs is reduced to about 1/6 instead of 0.53. The explained remaining variance is only 0.14 instead of 0.24, revealing that the selection incentive is dramatically lower when the risk adjusters are included in

the health-based risk adjustment model.

TABLE XIII: REMAINING SELECTION INCENTIVES WITH THE PROSPECTIVE MODELS

	Actual		Beck	
Ambul. Costs	0.881	**	0.795	**
Hospit. costs	0.533	**	0.352	**
Constant	-2,743	**	-2,344	**
N Obs	363,353		363,353	
R-squared	0.24		0.17	

	One Year - DRG		Two Year - DRG		Three Year - DRG		Four Year - DRG	
Ambul. Costs	0.776	**	0.728	**	0.683	**	0.691	**
Hospit. costs	0.176	**	0.155	**	0.149	**	0.123	**
Constant	-2,113	**	-2,071	**	-2,050	**	-2,109	**
N Obs	363,353		255,313		157,493		73,326	
R-squared	0.14		0.12		0.14		0.19	

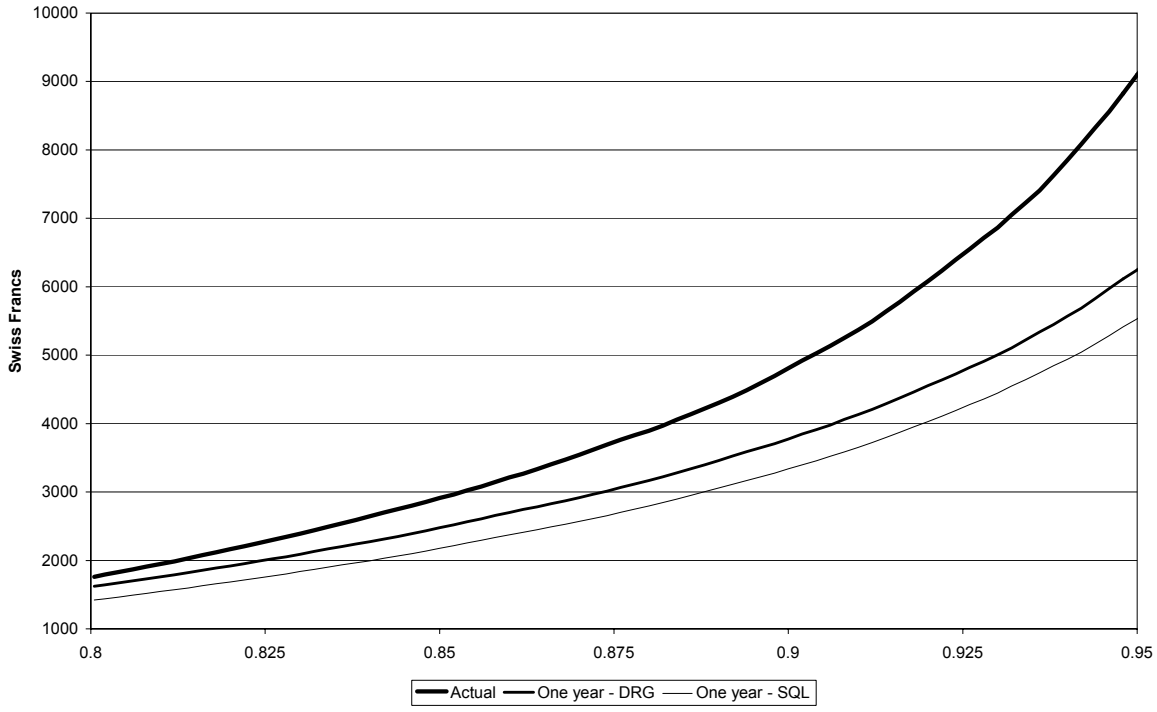
	One Year - SQL		Two Year - SQL		Three Year - SQL		Four Year - SQL	
Ambul. Costs	0.680	**	0.608	**	0.558	**	0.576	**
Hospit. costs	0.165	**	0.151	**	0.145	**	0.112	**
Constant	-1,863	**	-1,754	**	-1,700	**	-1,770	**
N Obs	363,353		255,313		157,493		73,326	
R-squared	0.11		0.09		0.11		0.14	

* significant at 5%; ** significant at 1%

The regression may not capture the correct impact. The goal of the risk adjustment model is a sizeable reduction in the expected gains of actively selecting out bad risks. We should therefore look more closely at the right tail of the distribution, and examine the impact on those identified as bad risks. We perform the following exercise: assume that the insurance firm uses one of the models in Table XIII to predict in which cases there will be a risk adjustment which is insufficient to covert the actual expenses in year 2002, that is a prediction is made using the above coefficients and current expenditures in hospital ad ambulatory in year 2001. Those cases for which the prediction is above a certain threshold are selected out. We do not focus on any specific value for this threshold, as we do not know how much the insurers should pay per

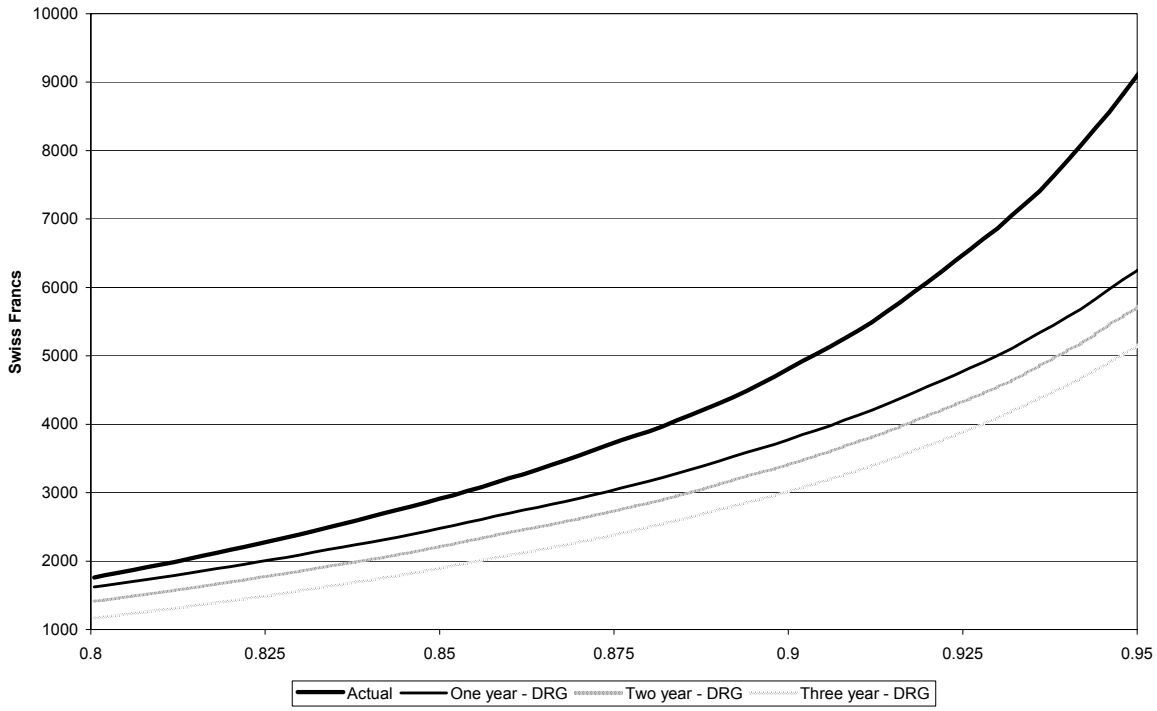
successfully repelled bad risk and how much they value the risk may damage their public image. The results of this exercise are presented in Figure IV and Figure V.

FIGURE IV: RANKED DIFFERENCES BETWEEN EXPECTED AND OBSERVED 2002 TOTAL EXPENDITURES- AP-DRG VS SQLAPE



We interpret the curves presented in Figure IV by saying that, for a threshold of CHF 2,000 in 2002 for example, the active selection would occur for 18.7.9% of the policyholders, whereas with the AP-DRG risk adjuster this percentage drops to 17.5., and to 16.0 in the SQLape case. The same values for a threshold of CHF 5,000 are 9.0%, 7.0% and 5.9% respectively. The graph presented in Figure V shows the progressive impact of the introduction of the medical categories risk adjusters. The selection percentages for the CHF 2,000 threshold are 17.5%, 16.1% and 14.4% for the one, two and three year models respectively. The same percentages for the CHF 5,000 are 7.0%, 6.1% and 5.2%.

FIGURE V: RANKED DIFFERENCES BETWEEN EXPECTED AND OBSERVED 2002 TOTAL EXPENDITURES – ONE-YEAR VS MULTI-YEAR



Finally we compare all the models we regard to their predictive power out of sample. The results are presented in Table XIV.

TABLE XIV: OUT OF SAMPLE R2- MULTI-YEAR MODELS

	Out of sample R2 DRG	Out of sample R2 SQLape
Actual	.27	.27
Insurance	.56	.56
Beck Type	.31	.31
One year model	.36	.40
Two year model	.38	.41
Three year model	.39	.42
One year two-part model	.36	.42
One year Tobit model	.34	.38
One year Heckman selection model	.36	.40

One can see significant improvements of the models which use medical information as compared to those which don't. There is a marginal increase from the one-year to the multi-year case and the SQLape classification performs better than the AP-DRG. But also in the SQLape case there is only a marginal improvement by using information from previous years. This suggests that careful construction of disease groupings may pay off in terms of health expenditure prediction, especially groupings that optimise the use of information from multiple years.

9 IMPLEMENTATION OF THE PROPOSED HEALTH-BASED RISK ADJUSTMENT MODEL

9.1 GENERAL PRINCIPLES

Based on the procedure used in this study, we observe that the general outline for the implementation of the risk adjustment model is quite simple: at a given time, all the hospitals in the system transmit their medical statistics for the previous year to the regulator, and simultaneously the sickness insurance funds transmit their past expenditure data on their policyholders. Then, the regulator merges the two data sets and performs three procedures:

1. Compute the mean cost for each risk group; these values are then defined as being the risk-adjusted payments for the next year;
2. Determine the list of applicable risk adjusters for each patient;
3. For each sickness insurance fund, apply the values obtained in the first step for their policyholder, and compute the sum.

Finally, they inform the sickness insurance funds of how much each of them should pay or receive from the risk adjustment fund.

These operations are very similar to those that are presently applied in Switzerland, except that additional medical information is used.

The results presented in this study have shown the interest and the feasibility of a prospective health-based risk adjustment model, based on one-year or multi-year medical information on prior hospitalisation. If such a health-based risk adjustment model is accepted in the future, the first step would be to ask all sickness insurance funds to send their 1998 to 2003 data to the Swiss Federal Statistical Offfice (SFSO/BFS): year of insurance, insured anonymous identifier (based on the same technology as for medical hospital data), sickness insurance fund identifier, ambulatory and hospital expenditures (with and without deductibles), birth year, gender, presence of a hospital stay, presence of a

nursing home stay, and contractual deductible. These data will then be matched with medical hospital records and, ensuring anonymity, transmitted in an anonymous way to the Swiss institution in charge of managing the risk adjustment mechanism.

The choice of the best patient classification system can be discussed further: the only condition is that the retained patient classification system is compatible with medical hospital nomenclatures (CIM-10 for diagnoses and CHOP-IV for procedures). This leaves some time to improve upon the medical adjustment scheme (see the "Future Research" Section).

Although the principle is quite simple, there are a few practical issues to discuss in view of the implementation of our type of risk adjustment model.

9.2 PRACTICAL ISSUES

9.2.1 Patients' identification (ID)

In order to guarantee that the information from hospitals is correctly matched with the information from sickness insurance funds, each individual should have an identification number. It should appear on the databases of both the sickness insurance funds and hospitals.

This number should be as anonymous as possible. Even if the database remains with the regulator, the risk of being able to trace patients should be minimised. The same anonymous linking code provided by the Swiss Federal Statistical Office (SFSO/BFS) should be used [OFS, 1998]. This procedure, based on a phonetic analysis of the surname and first name ("Soundex method"), the birth day and the gender, should provide an efficient mean to link different data bases (Sickness funds data and Hospital medical data, see Appendix C) without any risk to identify a specific patient. The simultaneous use of an independent anonymous identifier provided by the sickness funds could be useful to solve the issue of surname changing (marriage for instance). A test should be done to measure the rate of correct linkage, since the validation of the SFSO/BFS linkage procedure has been done only on a hospitalised population. In case of residual problems, the sickness funds identifier could be added to the

Hospital Medical data set to solve possible error in collecting surnames or first names.

9.2.2 Hospital failure to deliver the data in time

If we base our assumption on current practice, hospitals tend to take too long to deliver good quality databases. This problem seemingly hinders the project. Obviously, without complete data, the computation involved in the risk adjustment procedure may be completely wrong.

Two solutions are quite easy to implement as long as the number of hospitals failing to deliver such databases is small:

a) the computation of the risk-adjusted payments based only on the available information;

b) the computation of the risk-adjusted payments based on past information for the missing values.

Once the correct values are available, the correct amount that should have been paid to the sickness insurance funds and the difference between these and the amounts actually paid should be computed. This difference is kept as a credit or a debit for the next year's payments. Another solution is to compute the patients in that hospital as a risk group by itself, and compute the mean future cost based on the previous year's expenditures from inpatients two years ago in that hospital.

9.2.3 Hospital failure to deliver the correct database.

If, regardless of the length of time taken, the hospitals are not able to send the correct database, the problem then is quite different. This situation can occur if there is bad quality coding or a suspicion of fraud.

Two different approaches should be taken:

a) An investigation in the given hospital to determine the real reason behind the poor quality of the database. The main goal is to ensure that the following year's database will be better. Since hospitals do not benefit from the risk adjustment scheme, the only pressure is to make it obligatory, as it is

already today.

b) For this year only, consider the wrongly coded patients or the whole hospital as a risk group by itself.

9.2.4 Sickness insurance funds' failure to deliver the correct database in time

This situation should rarely arise if the sickness insurance funds are penalized by having their risk adjustment compensation withheld, while still having to pay a contribution to the risk adjustment fund.

9.2.5 Inpatient stays over New Year's Eve

This has much to do with the correct timing of the whole process. We are not able yet to determine the exact date at which we may reasonably obtain the complete databases from both sides every year.

In fact, we are facing a trade-off problem: on one hand we need data quickly to compute and inform the sickness insurance funds in good time of the risk adjustment payments they receive or have to pay; on the other hand, we need good quality data to compute accurate risk adjustment payment values. Once the system performs well, quality is important only in terms of categorising individuals according to the correct risk group, the computation of average excess risk being stable in the long run.

9.2.6 Severe illnesses which do not necessitate an inpatient stay.

During the initial phase, a complementary procedure should be available to assign a patient to a high-risk group, even if he does not require hospitalisation. This has still to be correctly defined, but a simple procedure could be to look at the medical information for those patients who have abnormally high ambulatory care without any inpatient stays (list of people on dialysis for instance).

9.2.7 Treatment of cantons

Health systems differ in every canton, leading to different levels of funding participation among the sickness insurance funds; the level is set at least 50%. This issue can be solved by fixing cantonal parameters. During the initial phase, a risk adjustment scheme may be computed independently for each canton, as long as discrepancies are not so high that it cannot be correctly computed for the overall system. However, it should be noted that the accuracy of the system is highly correlated to the size of the database, particularly for rare diseases.

9.2.8 Underestimation of risks

If health expenditures do rise every year, a systematic underestimation bias will occur in the computation of risk adjusters. Therefore sickness insurance funds may fear that risk adjustment will always be lower than real expenditures, and therefore they will lose money.

Since the adjustment payments are computed in such a way that the total money paid is equal to the amounts received, this fear may not exist for the sickness insurance funds as a whole.

9.2.9 Medical innovation

Medical innovation may completely change the outcome for some illnesses. For instance, fatal diseases for which there is currently no treatment available may be treated in the future using costly treatment; a new treatment may inversely cure illnesses which currently imply very costly and lengthy treatment. This observation implies that the *risk group definitions should be reviewed periodically*, not only the value of the risk-adjusted payments for the risk groups. With the help of the database, more subtle evolutions in treatment practice may be evaluated and lead to the redefinition of risk groups based on diagnostic information.

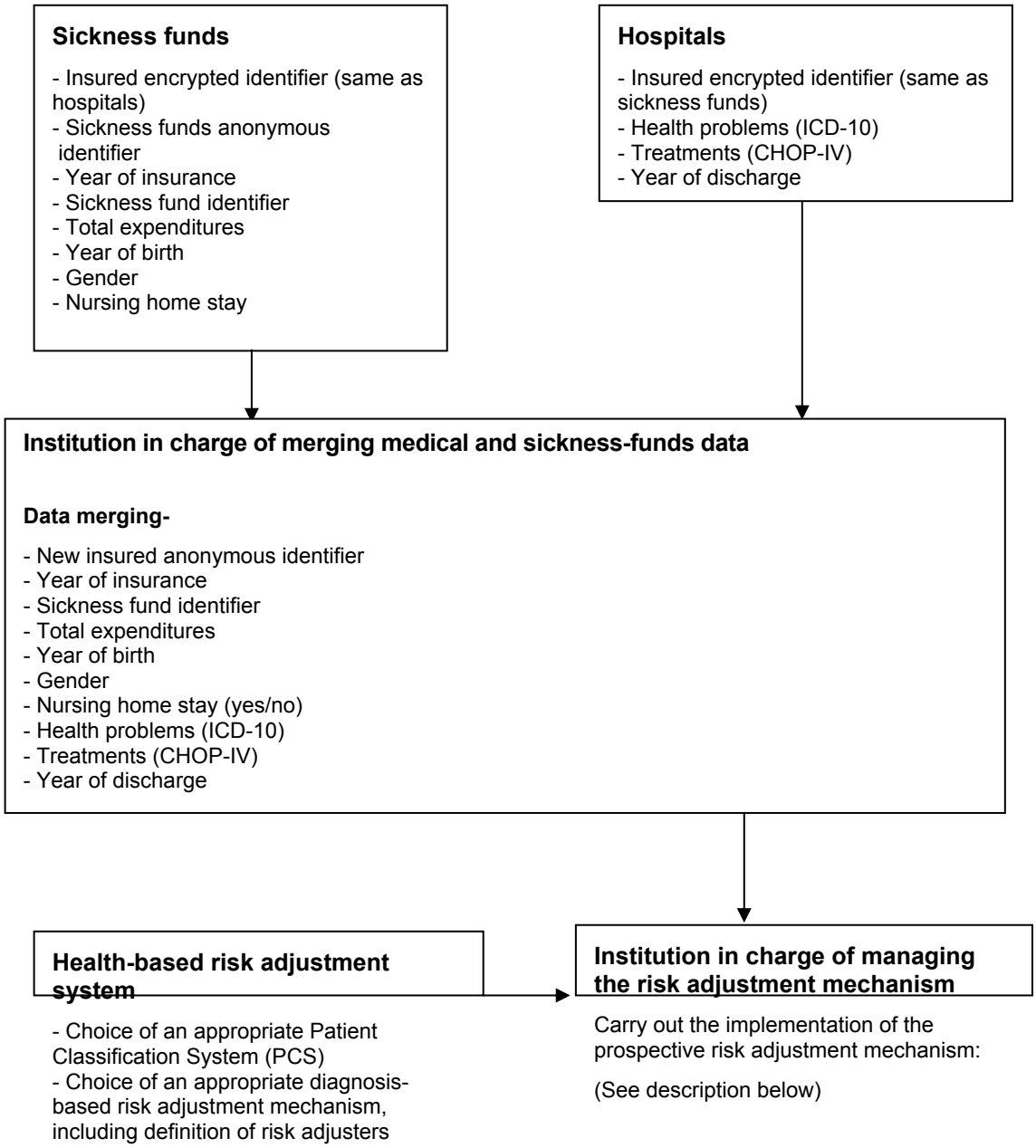
9.2.10 Responsibility of the institutions: hospitals, SFSO and the institution in charge of managing the risk adjustment mechanism.

The main work required in the implementation of the risk adjustment model is the annual construction of a valid database. Medical group assignment is strictly based on the same routinely available hospital data; either AP-DRGs or SQLape classification systems can be used without differences in the cost of information.

At the moment, hospitals send the medical observations to SFSO, which is responsible for developing the statistics. The institution in charge of managing the risk adjustment mechanism receives the data from the sickness insurance funds in order to compute the present risk-adjusted payments. Perhaps, these data should also be simultaneously transmitted to SFSO, using the same protected policyholder/patient ID.

No matter which institution is finally put in charge of managing the risk adjustment mechanism, the burden will not be much higher than today. Hospitals and the sick insurance funds still transmit a yearly database to this institution. The management of the new risk adjustment mechanism will not be different from the present procedure implemented by the institution in charge of managing the risk adjustment mechanism; only the amounts of risk-adjusted payments will be computed differently.

FIGURE VI: DATA PROCESSING FOR PROSPECTIVE RISK ADJUSTMENT



The following is a description of the prospective risk adjustment procedures, from a practical point of view, than can be carried out by the institution in charge of managing the risk adjustment mechanism:

- As soon as medical information from the year (t-1) is available, it performs in year (t) a regression of the expenditures of individual (i) in year (t-1), on the risk adjusters of individual (i) in period (t-2) (medical and demographic information, etc).
- With the estimated regression coefficients it can predict expenditures in year (t) for each insured individual in year (t) using medical information from year (t-1).
- We define the risk adjustment payment an insurer will receive (pay) for an insured individual, for any risk adjuster, to be the difference between predicted expenditure for year (t) for this risk adjuster and average predicted expenditure for year (t).
- The institution informs the insurers of the total he will receive or pay for all its enrolees;
- Note that this definition has at least two advantages in practical terms:
    - The computation of the risk adjustment payment can be done in year (t) as soon as information from the year (t-1) is available. This means that insurers are informed about what they are going to pay or receive quite soon. In fact, they can get this information before they set-up the premium for the following year. This is an important improvement over the retrospective system;
    - Like in the retrospective scheme, the sum of all risk adjustment payments is zero by definition.

We would like to draw the attention to an essential point of the mechanism which we propose. The link between the various data sources is obtained by using an encryption coding procedure which makes it possible to generate an anonymous code of connection. It is impossible to infer from this anonymous code any the variables having been used to generate it (first names, names, sex

and date of birth of the patients). This information, in a form which preserves the strict anonymity of the policyholders, is held by the institutions in charge of connecting the medical data bases and those of the insurers, and in charge of implementing the risk adjustment procedure. Never at any time have the insurers access to individual information coming from the medical records of the hospitals. The strict data confidentiality is thus guaranteed. Consequently, the risk adjustment mechanism which we propose respects the data protection of the policyholders. As a result, this mechanism does not provide under any circumstances additional information on the health status of the policyholders that the insurers could use in a strategic way on the level of the basic insurance as well as on the level of the supplemental insurance.

10 FUTURE RESEARCH

In the development of the health-based risk adjustment models presented in our study, we frequently mentioned a number of areas of research which are worth exploring further. In this section, we summarise those which appear to be the most important and the most pressing.

First, in our discussion of the risk adjustment payment based on the mapping of the AP-DRGs in the medical adjustment categories, we often pointed out two limitations of this particular PCS: a) they do not explicitly adjust for severity of illness and b), the AP-DRGs assignment rule rests on both patient diagnoses and procedures. Yet, risk adjustment models should exclude treatments or procedures as far as possible, since the classification is more difficult due to the underlying illness not being known. Thus, we would like to explore as soon as possible the use of other types of PCS which explicitly take into account the severity of illness and which are mainly diagnosis-based. This is because the rationale for prospective risk adjustment models is different than for the funding of hospitals; it seeks to predict next year's health care expenditures using the current year's expenditure, mostly for chronic illness.

Second, we noted the importance of previous year outpatient expenditures, together with inpatient expenditures, to predict future health care expenditures by sickness insurance funds. We did not use these expenditures as risk adjusters since they create perverse incentives. However, we used a PCS for hospital encounters as a substitute for inpatient expenditures. The next step would consist in extending the medical information to outpatient services. This is not feasible in Switzerland presently, as no reliable patient classification system for outpatient services is available. Therefore, we intend to set up a collaborative project, notably with federal institutions, to develop a patient classification system for outpatient services, as described in Section 3.3.2. At the same time, we also intend to set up another collaborative program to use information on chronic conditions deduced from the prior use of prescribed drugs together with outpatient care diagnoses to improve the risk adjustment

scheme based on health status information. This is because, as already mentioned in Section 3.3.2, outpatient care diagnoses and information on chronic conditions deduced from the prior use of prescribed drugs have been shown to be good predictors of future health care expenditures. The purpose of this would be to develop one-year and multi-year, health-based risk adjustment models, using diagnostic information on the previous use of either inpatient or outpatient services, or of prescribed drugs, which could be mentioned in individual patient encounters. It is expected that this full health-based risk adjustment model would improve the predictive accuracy of the risk adjustment model, presented in our study, which uses medical information from prior hospitalisation only.

Third, we would like to analyse the role of deductibles as a possible risk selection tool used by sickness insurance funds. In fact, a recent study by Geoffard et al. (2003) has shown that there is an adverse selection effect due to the choice of deductibles in the health insurance market in Switzerland. Thus, a possible strategy of sickness insurance funds would be to risk select through a marketing policy discouraging possible candidates to choose low deductibles. If this turns out to be the case, then it would be natural to include the high deductibles among the risk adjusters.

Finally, and despite all these efforts to devise a good risk adjustment model and improve its predictive accuracy, we may still underestimate very high risks. This is why we intend to explore the possibility of setting up an arrangement in the shape of a combination of a prospective risk adjustment model with retrospective risk sharing. The main two forms of risk sharing we have in mind are stop loss and “risk sharing for high-risks” types of reinsurance mechanisms [Van de Ven and Van Vliet (1992), Van Barneveld et al. (1996)].

Needless to say that it would be impossible to carry on with these future directions of research without the active collaboration of sickness funds and health care providers. We shall make every effort to consolidate our relationships with the institutions which helped us to construct the database used this study, and to extend this relationship to other institutions.

11 CONCLUSION

In this study we have developed health-based prospective risk adjustment models for Switzerland. The primary reason for implementing risk adjustment is to correct for risk selection problems and to prevent "cream skimming". This selection may occur because community rating implies predictable profits on low-risk consumers and predictable losses on high-risk consumers, and thus sickness insurance funds have an incentive to avoid bad risks and appeal to good risks. Furthermore, as we have illustrated with our data, this incentive is reinforced by the fact that health expenditures are very highly concentrated in relatively few individuals.

As we already pointed out in the Introduction, it is clear from this concentration of expenditures that health status is an important factor for predicting variations in annual per-person health care expenditures in risk adjustment models. There are in fact different ways to develop health-based risk adjustment models, according to the nature of the variables which are used to describe individual health status. The identification of individuals among whom expenditures are very highly concentrated is crucial to the effectiveness of risk adjustment.

Our econometric results have broadly confirmed that the present risk adjustment model is inadequate; it does not fully provide incentives to prevent cream skimming. Indeed, apart from the fact that it is a retrospective model, it does not use health-related risk adjusters; it assumes that health care costs are only correlated with the two variables "age" and "gender", and therefore does not properly take into account, for example, young people with very costly illnesses or, the elderly in good health. We observed that the predictive power of this "demographic variable" is extremely poor and thus, leaves a lot of room for risk selection by sickness insurance funds.

Beck (1999) has suggested a model which accounts for the health status of the insured thanks to the introduction of a dichotomous variable ("yes" and "no") for prior hospitalisation as an additional risk adjuster to the formula currently in

use. The formula suggested by Beck (1999) also changes the method of risk adjustment calculation from retrospective, as currently in use, to prospective.

The present study attempts to go beyond the very important refinement of the existing risk adjustment model proposed by Beck (1999). Its aim is to develop a prospective risk adjustment model at the cantonal level which introduces medical information from prior hospitalisation as a risk adjuster, jointly with age and gender.

We consider in our study two types of risk adjustment models, which we called "one-year model" and "multi-year model". The one-year model consists of predicting expenditures (i.e. sickness insurance funds' payments to providers) in the current year (t) for the next year (t+1) with data on hospitalisation (gender, age, hospital and nursing home stays, medical adjustment categories) available at t. As indicated earlier, the medical adjustment categories are obtained alternatively from AP-DRGs and SQLape classifications.

We also considered the multi-year risk adjustment model which consists of predicting sickness insurance funds expenditures in the current year (t) for the next year (t+1) with available data on hospitalisation from year t and the previous years (t-1, t-2, etc. with a maximum of four years in our study) [Lamers and van Vliet (1996) and Lamers (1997)]. The multi-year models are particularly useful from a risk adjustment perspective, as most chronic illnesses can be cumulated over many years, i.e. the condition is considered as present if it occurred during any previous year, but some conditions do not have a "multi-year" effect. To this end, we have constructed what we called a set of "aggregation rules" to improve identification of the risk adjustment groups on the basis of medical information from prior hospitalisation from previous years.

In order to evaluate the health-based risk adjustment model that we have developed in this study, we shall use two criteria suggested by Van de Ven and Ellis (2000), namely *feasibility* and *appropriateness of incentives*.

11.1 The Feasibility criteria

We quote from Van de Ven and Ellis (2000): "Administrative *feasibility* is the requirement that the measures are feasible to obtain for all potential enrolees

without undue expenditures of time or money. Information that is routinely collected, standardized and comparable across different health plans, and measures that are easily validated have greater feasibility than measures that require separate data collection, validation and processing."

It is clear from the description of the risk adjustment procedure outlined in Section 9 that it perfectly satisfies the administrative feasibility criteria. Indeed, the general outline for the implementation of the risk adjustment model is quite simple: at a given time, all the hospitals in the system transmit their medical statistics from the previous year to the regulator, and the sickness insurance funds transmit their past expenditure data for their policyholders. The regulator then merges the two data sets. Once the data set is constructed, the regulator carries out some simple computations which are very similar to those that are applied in the present risk adjustment model in Switzerland, except that additional medical information is used. In Section 9 we also provide some solutions to a few practical issues to guarantee the administrative feasibility of the proposed health-based risk adjustment model.

11.2 THE APPROPRIATENESS OF INCENTIVE CRITERIA

The most important criterion for evaluating risk adjustment models is the extent to which they reduce sickness insurance funds' incentives to select good risks [Van de Ven and Ellis (2000)].

It is clear from the results presented in Section 7 and Section 8 that the prospective health-based risk adjustment models that we have developed in our study dramatically improve upon the demographic risk adjustment model. First, the effect of age structure is significantly reduced, and this result shows that we could significantly improve the risk adjustment payment scheme for those who did not have any inpatient stays during the current year. Second, we were able to compute risk adjustment payments according to medical risk category. Third, our computations show that the health-based risk adjustment models considerably reduce the incentive to cream skim through much better risk adjustment payment systems.

These models also significantly improve upon the type of risk adjustment model proposed by Beck (1998, 1999). The results highlight the great importance of, and urgent need for an appropriate health-based risk adjustment model using medical information from prior hospitalisation.

It is important to observe that the "nursing home" variable is highly significant. We thus found it necessary to include an explicit, separate risk adjustment payment for this variable. The purpose of this risk adjuster is to ensure that the risk adjustment model takes into account the average expenditures of the insured population resident in a registered nursing home.

It is also illuminating to compare the risk adjustment payments derived from the SQLape and AP-DRGs mappings. As mentioned earlier, the purpose of using the SQLape in our study is to illustrate the potential improvement that could be achieved from alternative classification systems to AP-DRGs which explicitly take into account the severity of illness. As we have shown in Section 7, the risk adjustment payment is much smaller for a normal delivery, whereas it is almost twice the amount of the risk adjustment payment for a patient with terminal insufficiency (DED). This comparison shows the importance of the consideration of the severity of illness in the development of a prospective health-based risk adjustment model.

Last but not the least, the results for the multi-year health-based risk adjustment models presented in Section 8 convincingly illustrate the usefulness of this type of model. By using available data on prior hospitalisation from previous years (t-1, t-2, etc.), one can better identify over time the risk adjustment groups, notably for most chronic illnesses which can be cumulated over many years. Here again, the comparison between AP-DRG and SQLape mappings is quite informative; by taking into account the severity of illness, SQLape better identifies the extreme bad risks than the AP-DRG. This is a crucial result as far as the effectiveness of risk adjustment is concerned, since it allows a better identification of the individuals among whom expenditures are very highly concentrated.

These comparisons provide an additional strong incentive to test as soon as possible other patient classification systems which explicitly take into account the severity of illness, and which are mainly diagnosis-based and exclude treatments or procedures where possible.

To conclude, however, we would like to emphasise that we have achieved our main objective, which is to demonstrate, not only that the development of a prospective health-based risk adjustment model in Switzerland using medical information from prior hospitalisation is administratively feasible, but would also considerably reduce the incentive to select risks through much better risk adjustment payment systems.

The health-based risk adjustment type of model suggested in our study is substantially more effective than the model currently in use, which looks only at demographic and geographic factors.

12 REFERENCES

Arrow, K., and G. Debreu (1954), "Existence of an equilibrium for a competitive economy," *Econometrica* 22: 265-290.

Ash A, Porells F, Gruenberg L et al. (1989), "Adjusting medicare capitation payments using prior hospitalisation date," *Health Care Financing Review*, 10: 17--29.

Ash A, Ellis R. P., Yu W. et al. (1998), *Risk adjusted payment models for the non-elderly.* Final report, for Health Care Financing Administration, June.

Averill RF, Muldoon JH, Vertrees JC, Goldfield NI, Mullin RL, Fineran EC, Zhang BS, Steinbeck B, Grant T. (1998), "The evolution of casemix measurement using Diagnosis Related Groups (DRGs)," *3M HIS Working Paper*, 5.

Averill, R. F., Goldfeld, N. I., Eisenhandler J. et al. (1999), *Development and Evaluation of Clinical Risk Groups (CRGs)*, 3M Health Information Systems, Wallingford, CT

Antioch K. M., Walsh M. K. (2000), "Funding issues in Victorian Hospitals: the risk-adjusted vision beyond casemix funding," *Aust. Health Rev.*, 23, 145—153

Antioch K. M., Walsh M. K. (2002), "Risk-adjusted capitation funding models for chronic disease in Australia: alternatives to casemix funding," *European Journal of Health Economics*, 3, 83—93

Antioch KM and Walsh MK (2003) "The risk adjusted vision beyond casemix (DRG) funding in Australia: International lessons in high complexity and capitation. *European Journal of Health Economics*. Published online 10 December, 2003.

Beck, K. (1998), "La compensation des risques – fondement d'une concurrence appropriée dans l'assurance-maladie" (Translation of the title: Risk adjustment -- basis for an appropriate competition in health insurance), CSS Insurance, Lucerne.

Beck, K. (1999), "Risk adjustment and competition in Swiss health insurance market – developments, shortcomings and alternatives." *Mimeo*.

Beck, K., Spycher S., Holly, A. and Gardiol, L. (2001), "Risk adjustment in Switzerland," Mimeo, August 2001. To appear in *Health Policy*

Beck, K. and Zweifel P. (1998), "Cream-skimming in deregulated social health insurance: evidence from Switzerland ," in *Health, the Medical Profession and Regulation*, P. Zweifel (ed.), Kluwer Academic Publishers, Dordrecht.

Carter, G.M., Bell, R.M., Dubois, R.W. et al. (1997) *A clinically detailed risk information system for cost.* Report for HCFA, RAND report number DRU-1731-1-HCFA November.

Clark, D.O., von Korff, M., Saunders, K., Baluch, W.M. and Simon, G.E. (1995) "A chronic disease score with empirically derived weights," *Medical Care,* 33:783-795.

Dormont, B., Holly A., Milcent C., and Yalcin T. (2003), "Les systèmes de classification des patients: une brève revue de la littérature,", Final Report to the Swiss Federal Office of Social Insurance, Institute of Health Economics and Management (IEMS), Lausanne, Switzerland

Duan, Naihua, Willard G. Manning, Carl N. Morris and Joseph P. Newhouse (1983), "A comparison of alternative models for the demand for medical care", *Journal of Business & Economic Statistics*, Vol. 1, No. 2, pp. 115—126.

Dunn, D. L., Rosenblatt, A., Tiaira, D. A. et al. (1996) *A comparative analysis of methods of health risk assessment*. Final report to the Society of Actuaries.

Dunn, D. L. (1998), "Applications of Health Risk Adjustment: What Can Be Learned from Experience to Date ?," *Inquiry,* 35

Ellis, R.P. and Ash, A. (1995), "Refinements to the Diagnostic Cost Group model," *Inquiry*, Winter:418-429.

Ellis, R. P. and McGuire T. G. (1986), "Providers behavior under prospective reimbursement: cost sharing and supply," *Journal of Health Economics*, 5: 129-151.

Ellis, R. P. and McGuire T. G. (1993), "Supply-Side and Demand-Side Cost Sharing in Health Care," *Journal of Economic Perspectives,* 7: 135-151.

Ellis, R. P., Pope G. C., Iezzoni L. I., Ayanian J. Z., Bates, D. W., Burstin H., and Ash A. S. (1996a), "Diagnosis-based risk adjustment for Medicare capitation payments," *Health Care Financing Review*, 12: 101-128

Ellis, R. P., Pope G. C., Iezzoni L. I. et al. (1996b), *Diagnostic Cost Group (DCG) and Hierarchical Coexisting Conditions and Procedures (HCCP) models for medicare risk adjustment*. Final Report, Health Economics Research, for Health Care Financing Administration, April.

Epstein, A. M. and Cumella E. J. (1988), "Capitation payment: using predictors of medical utilization to adjust rates," *Health Care Financing Review*, 10: 51--69.

Evers S, Voss G, Nieman F, Ament A, Groot T, Lodder J, Boreas A, Blaauw G (2002). "Predicting the cost of hospital stay for stroke patients: the use of diagnosis related groups," *Health Policy* 61; 21-42.

Fetter R.B., Shin Y., Freeman J.L. (1980). Casemix definition by DRGs. *Medical Care*, 18, 1-53 (supplement).

Fetter, RB (ed) (1991): *DRGs: Their Design and Development*. Ann Arbor, MI: Health Administration Press.

Folland S., Goodman A. C., Stano M. (2001), *The economics of health and health care,* 2nd ed., Prentice-Hall, Inc.

Geoffard, P.-Y., Grandchamp Ch. and Gardiol L. (2003), "Separating selection and incentive effects: an econometric study of Swiss health insurance

claims data," *mimeo*.

Giacomi, M., Luft H. S., and Robinson J. C. (1995), "Risk adjusting community rated health plan premiums: a survey of risk assessment literature and policy applications," *Annual Review of Public Health*, 16: 401 – 430.

Guesnerie, R. (1995), "The genealogy of modern theoretical public economics: From first best to second best," *European Economic Review* 39: 353-381.

Holly A, Gardiol L., Domenighetti G. and Bisig B. (1998), "An Econometric Model of Health Care Utilization and Health Insurance in Switzerland," in Papers and Proceedings of the Twelfth Annual Congress of the European Economic Association, edited by Philippe Jehiel, *European Economic Review*, 42: 513-522.

Holly, A. and Pentsak, Y. (2003): "Maximum likelihood estimation of the conditional mean E(y/x) for skewed dependent variables in four-parameters families of distributions", Institute of Health Economics and Management (IEMS), *Mimeo*

Hornbrook, M.C., Goodman, M.J. and Bennett, M.D. (1991), "Assessing health plan case mix in employed populations: ambulatory morbidity and prescribed drug models," in R.M. Scheffer and L.F. Rossiter, eds., *Advances in health economics and health services research*, Vol. 12. JAI Press, Greenwich, 197-232.

Hornbrook, M.C. and Goodman, M.J. (1996), "Chronic disease, functional health status, and demographics: a multi-dimensional approach to risk adjustment," *Health Services Research*, 31:283-307.

Jackson T (2001), "Using computerized patient-level costing data for setting DRG weights: the Victorian (Australia) cost weight studies," *Health Policy*;56:149-163.

Keeler, E.B., Carter, G.M. and Trude, S. (1988), "Insurance aspects of DRG outlier payments," *Journal of Health Economics*, 7:193-214.

Keeler, E.B., Carter, G.M. and Newhouse, J.P. (1998), "A model of the impact of reimbursement schemes on health plan choice," *Journal of Health Economics,* 17:297-320.

Kroneman M, Nagy J (2001), "Introducing DRG-based financing in Hungary: a study into the relationship between supply of hospital beds and use of these beds under changing institutional circumstances." *Health Policy*; 55(1):19-36.

Kronick R, Dreyfus T, Lee L, Zhou Z. (1996), "Diagnostic risk adjustment for Medicaid: the disability payment system," *Health Care Financing Review*, 17: 7--34.

Lamers, L. M. (1997), *Capitation payments to competing Dutch sickness insurance funds based on diagnostic information from prior hospitalizations*, Ridderprint, Ridderkerk.

Lamers, L.M. (1998), "Risk adjusted capitation payments: developing a diagnostic cost groups classification for the Dutch situation," *Health Policy,* 45:15-32.

Lamers, L.M. (1999), "Pharmacy costs groups: a risk adjuster for capitation payments based on the use of prescribed drugs?", *Medical Care*, 37: 824 – 830.

Lamers, L. M. and van Vliet R. C. J. A. (1996): "Multiyear diagnosis information from prior hospitalisation as a risk-adjuster for capitation payments," *Medical Care*, 34, pp 549—561.

Lamers, L. M. and van Vliet, R. C. J. A., (2003), "Health-based risk adjustment. Improving the pharmacy-based cost group model to reduce gaming possibilities," *European Journal of Health Economics*, 4: 107 – 114.

Manning, Willard G., N. Duan and W. H. Rogers (1987), "Monte Carlo evidence on the choice between sample selection and two-part models", *Journal of Econometrics*, 15, pp. 59—82.

Manning, W. G. (1998), "The logged dependent variable, heteroskedasticity, and the retransformation problem," *Journal of Health Economics*, 17: 283-296.

Manning, Willard G. and John Mullahy (2001), "Estimating log models: to transform or not to transform?", *Journal of Health Economics*, 20, pp. 461—494.

Manning, W. G., Basu A. and Mullahy J. (2003), "Generalized modelling approaches to risk adjustment of skewed outcomes data", *Mimeo*

Mullahy, J. (1998), "Much ado about two: reconsidering transformation and the two-part model in health econometrics," *Journal of Health Economics*, 17: 247-282.

NCHS (2002):ICD-9-CM The International Classification of Diseases, Ninth Revision, Clinical Modification. 6th Edition, NCHS

Newhouse, J. P. (1986), "Rate adjusters of Medicare capitation," *Health Care Financing Review*, Annual Supplement: 45-55.

Newhouse, J. P. (1991), "Pricing and Imperfections in the Medical Care Marketplace," in *Health Economics Worldwide*, eds. Peter Zweifel and H. E. Frech; Amsterdam: Kluwer Academic Publishers, 3-22.

Newhouse, J. P. (1994), "Patients at Risk: Health Reform and Risk Adjustment," *Health Affairs*, 13: 132-146.

Newhouse, J.P. (1996) "Reimbursing health plans and health providers: efficiency in production versus selection," *Journal of Economic Literature*, 34:1236-1263.

Newhouse, J. P. (1998), "Risk adjustment: where are we now ?," *Inquiry*, 35: 122-131.

Newhouse, J. P., Buntin M. B., and Chapman J. D. (1997), "Risk Adjustment and Medicare," *Health Affairs*, 16: 26-43.

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

Office fédéral de la statistique. Statistique médicale des hôpitaux, conception

détaillée 1997. Neuchâtel, OFS, 1998.

Pope, G.C., Adamache, K.W., Khandker, R.K. and Walsh, E.G. (1998a) "Evaluating Alternative Risk Adjusters for Medicare," *Health Care Financing Review*.

Pope, G., Ellis, R., Liu, C.F. , Ash, A., Iezzoni, L., Ayanian, J., Bates, D. and Burstin, H. (1998b) *Revised Diagnostic Cost Group (DCG)/Hierarchical Coexisting Conditions (HCC) Models for Medicare Risk Adjustment*. Final Report to HCFA, Health Economics Research, Waltham, MA, February.

Pope, G.C., Liu, C.F., Ellis, R.P., Ash, A. S., Iezzoni, L., Ayanian, J. Z., Bates, D.W., Burstin, H., Adamache, K. and Gilman, B. (1999*) Principal Inpatient Diagnostic Cost Group Models for Medicare Risk Adjustment*. Final Report to HCFA, Health Economics Research, Waltham MA, February.

Pope, G. C., Ellis R. P., Ash A., Liu C.-F., Ayanian J. Z., Bates D. W., Burstin H., Iezzoni L., Ingber M. (2000), "Principal Inpatient Diagnostic Cost Group Model for Medicare Risk Adjustment," *Health Care Financing Review*, 21, pp. 93--118

Rodrigues, J-M, Trombert Paviot, B., Martin, C. DRG information system, healthcare reforms and innovation of management in the Western countries during the 90s: where are the key success factors? *Casemix Quaterly*, 4(1), 2002

Selden, Th. M. (1990), "A Model of Capitation," *Journal of Health Economics,* 9: 397-409.

Spycher, S. (1999), "Wirkungsalalyse des Risikoausgleichs in der Krankenversicherung "(Translation of the title: Analysis of the impact of risk adjustment in social health insurance), Büro Bass, Bern.

Stiglitz, J. (1993), *Information and Economic Analysis*, Oxford University Press.

Thomas, J.W. and Lichtenstein, R. (1986), "Including health status in Medicare's adjusted average per capita cost capitation formula," *Medical Care*, 24:259-275.

3M Health Information Systems (1999), "Development and evaluation of Clinical Risk Groups (CRGs)," Final Report, April 30.

Van Barneveld, E.M., van Vliet, R.C.J.A. and van de Ven, W.P.M.M. (1996) "Mandatory high-risk pooling: an approach to reducing incentives for cream skimming", *Inquiry*, 33:133-143.

Van de Ven, Wynand and B. van Praag (1981), "Risk aversion and deducibles in private health insurance: Applications of an adjusted Tobit model to family health care expenditures", in *Health, Economics, and Health Economics*, eds. J. Van der Gaag and M. Perlman, Amsterdam: North Holland, pp. 125—147.

Van de Ven, W.P.M.M. and van Vliet, R.C.J.A. (1992) "How can we prevent cream skimming in a competitive health insurance market? The great challenge for the '90,s'" in P. Zweifel and H.E. French, eds., *Health Economics Worldwide*. Kluwer Academic Publishers, the Netherlands, 23-46.

Van de Ven, W. and Ellis, R. P. (2000), "Risk adjustment in competitive health plan markets," in A.J. Culyer and J.P. Newhouse (eds.) *Handbook of Health Economics*, Chapter 17, North Holland Publishing Company.

Van Vliet, R. C. J. A. and Lamers L. M. (1998), "The high costs of death: should health plans get higher payments when members die ?," *Medical Care*, 36: 1451 – 1460.

Weiner, J.P., Dobson S., Maxwell S., Coleman, Starfield B., Anderson (1996) "Risk adjusted Medicare capitation rates using ambulatory and inpatient diagnoses," *Health Care Financing Review*, 17:77-100.

Weiner, J.P., Starfield B., Steinwachs, D. and Mumford, L. (1991) "Development and application of a population oriented measure of ambulatory care case mix," *Medical Care*, 29:452-472.

WHO (1976): ICPM – International Classification of Procedures in Medicine, WHO

WHO (1997): ICD-10 The Tenth Revision of the International Statistical Classification of Diseases and Related Health Problems, volumes 1,2. WHO

Woolf Neville (1998). *Pathology. Basic and systemic*. London, Saunders.