ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 83/84 – 2006

Incentives for Risk Selection and Omitted Variables in the Risk Adjustment Formula*

Erik SCHOKKAERT†, Carine VAN DE VOORDE‡

ABSTRACT. – Risk adjustment in health insurance raises the question of how to treat variables which influence health care expenditures but do not capture acceptable costs differences. We argue that these variables should be included in the explanatory model and neutralized afterwards for the computation of the premium subsidies. This explicit approach is better than the conventional approach in removing the incentives for cream-skimming. We illustrate the empirical relevancy of the problem with data for Belgium.

Incitations à la sélection des risques et variables omises dans les facteurs d'ajustement du risque

RÉSUMÉ. – L'ajustement du risque dans l'assurance des soins de santé soulève une question quant au traitement des variables ayant un impact sur les dépenses en soins de santé qui ne capturent pas de différences de coûts socialement acceptables. Nous soutenons que ces variables devraient quand même être incluses dans le modèle explicatif, mais ne pas être prises en compte lors du calcul des subsides de compensation. Cette approche est meilleure que l'approche standard pour neutraliser les incitations à la sélection des profils de risque. Une application empirique sur des données belges illustre notre propos.

* We thank two anonymous referees, Robert Nuscheler and the participants of the Journées Louis-André Gérard-Varet on "Health, Insurance and Equity" in Marseille for their useful comments. The empirical example presented in the appendix results from joint work with Paul Kestens and Rodrigo Ruz-Torres (DULBEA, ULBruxelles). They are not responsible for the ideas presented in this paper, however.
† E. Schokkaert: KULeuven, Centre for Economic Studies, Naamsestraat 69, B-3000 Leuven (Belgium), email: erik.schokkaert@econ.kuleuven.be.
‡ C. Van de Voorde: KULeuven, Centre for Economic Studies, Naamsestraat 69, B-3000 Leuven (Belgium), email: carine.vandevoorde@econ.kuleuven.be.

1 Introduction

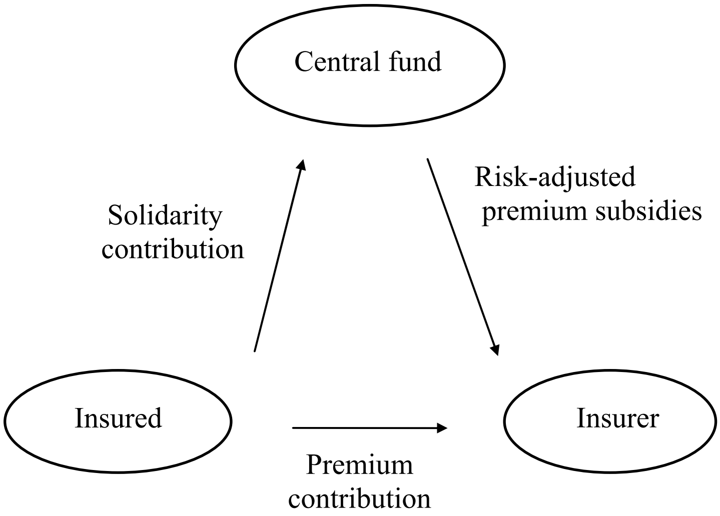
Many countries have introduced some risk adjustment into their system of health insurance. The basic rationale of risk adjustment within a system of managed competition can be easily explained with Figure 1.[1] Health insurance is compulsory but all citizens can enroll with their own preferred health insurer. They pay (possibly income-dependent) solidarity contributions to a central fund and premium contributions to the insurer. The purpose of the regulator is to increase equity by avoiding ethically undesirable differential treatment of different patient groups. More specifically, she imposes restrictions on the acceptable differentiation in premium contributions. In actual practice these are community rated.[2] Since health insurers have to reimburse health care expenditures, the imposition of community rating without further adjustment would create incentives for cream-skimming - another instance of undesirable differential treatment. This is precisely the reason why the central fund redistributes the solidarity contributions over the insurers through a system of risk-adjusted premium subsidies per member. With ideal risk adjustment, these premium subsidies exactly compensate for the differences in the risk, i.e. the expected expenditures of the individual members, and they therefore remove all incentives for risk selection. At the same time, they should not be based on the actual *ex post* expenditures, but should be given *ex ante*: in this way the insurers will remain motivated to curb (ex post) expenditures. More efficient insurers will have the opportunity to lower their (community rated) premium and hence make themselves more attractive for new members.

In all the countries where risk adjustment has been introduced, the actual formula is derived from an analysis of actual expenditures (see van de Ven *et al.*, 2003, for a comparison of the practice in different European countries). Risk adjusted premium subsidies are then calculated as the expenditures predicted from a regression model or as the relevant cell means. However, in principle it makes no sense to equate "acceptable" and "observed" expenses because this would destroy all incentives for efficiency. Risk-adjusted premium subsidies should be based not on actual but on "acceptable costs", i.e. "the costs which are generated in delivering a specified basic benefits package containing only medically necessary and cost-effective care" (Van de Ven and Ellis, 2000, p. 767). Or, to put it in a different way, "observed expenses are determined by many factors, not all of which need to be used for calculating the risk-adjusted subsidies" (Van de Ven and Ellis, 2000, p. 768). This raises the basic question: how to go from "observed" to "normative" expenditures?

To make this question more concrete, consider the large variety of factors explaining observed medical expenditures. The most important one is health status (morbidity); yet it is clear that at the level of the insured income, the level of own payments and subjective tastes also play a role. Moreover, health status is not only determined by what philosophers (Dworkin, 1981) have called "brute luck" but

1. Figure 1 describes the system in countries like the Netherlands, Belgium and Israel. In Switzerland and Germany there are no solidarity contributions but at the end of the period there is a transfer from sickness funds with a good risk profile to sickness funds with a bad risk profile. It is easy to show that the two systems are equivalent - see Van de Ven *et al.* (2003).
2. In Switzerland premiums are allowed to be different in different regions. In fact, one could argue that some premium differentiation can be equitable. We will return to this point in section 3.2.

FIGURE 1
Risk adjustment with a central fund



also by lifestyle variables (such as smoking and drinking habits) which can be controlled by the patient. In addition, provider behavior leads to differences in health care expenditures for individuals at the same morbidity level: (cost-ineffective) variations in medical practice may depend on the level of medical supply, on the price of medical care and on the way providers are remunerated.

No existing risk adjustment system in the world takes into account all these variables. In fact, there are basic reasons of principle not to include all of them - precisely because not all the resulting differences in health care expenditures are "acceptable". If interregional differences in medical supply lead to differences in provider behavior and therefore to differences in costs, this is an indication that not all these medical practices are cost-effective. If differences in health status follow from differences in lifestyle variables, equity perhaps does not necessarily require that the resulting expenditure differences are fully compensated. This suggests already two reasons for not including all variables in the calculation of the premium subsidies:

- some patient characteristics, that have an influence on medical expenditures, can be considered to be the patients own responsibility and, hence, offer no ground for compensation. Lifestyle variables are the most obvious example.
- variables that are directly linked to provider or insurer behavior should not be included in the premium subsidies. They reflect differences in efficiency and the whole idea of (prospective) risk adjustment is precisely to create incentives for insurers to act on these variables.

We call variables in either of these categories "illegitimate" risk adjusters or R-(responsibility) variables. Variables which should be included in the risk adjustment formula will be called "legitimate" risk adjusters or C-(compensation) variables.[3]

3. van de Ven and Ellis (2000) call the two sets "S"-type (for solidarity desired) and "N"-type (for solidarity not desired) respectively. We prefer the C-R terminology because it is also used in the social choice literature - see, e.g., the overview chapter of Fleurbaey and Maniquet (2003).

This raises immediately an ethical or political question: what variables should be seen as C-variables? Different options taken by different countries reflect different social and political values. The stronger is the aversion towards unequal treatment of different groups, the larger the set of C-variables. We will not treat this difficult choice problem in this paper but start from the assumption that the regulator in one way or another has taken a decision about what should be compensation variables. Once this choice has been made, however, a second (at first sight more technical) question arises. How to estimate the effect of the C-variables on the acceptable cost level? And more specifically, how to deal with illegitimate risk adjusters, i.e. variables that do have a significant effect on actual expenditures but should not be taken up in the risk adjustment formula? This question is the main focus of this paper.

In what could be called the "*conventional approach*" to risk adjustment, followed in the bulk of the empirical work and in actual policy applications, the effect of the illegitimate risk adjusters is largely neglected. This work focuses on the gradual introduction of better information on legitimate C-variables, health status being the most prominent example. Illegitimate R-variables are simply not included in the regression equations (or in the calculation of the cell means) used to compute the premium subsidies. The decision not to include R-variables mostly remains implicit. Sometimes it is made explicit, however. When diagnostic information or information about the use of pharmaceuticals is introduced in the risk adjustment formula, there is a keen awareness that this could create incentives for DRG-upcoding or for prescribing certain drugs when they are not really needed. Expenditures in previous years offer another obvious example; there is no doubt about their explanatory power but it is usually considered that introducing them in the risk adjustment formula would seriously impede the incentives for cost efficiency. This leads to a third category of illegitimate risk adjusters:

- variables which may contain useful information on the health status of patients when they are not used for the calculation of the premium subsidies, but which create incentives for manipulation or upcoding as soon as they are included in the risk adjustment formula.

As with the other illegitimate risk adjusters, the usual practice in conventional risk adjustment is to omit such variables not only from the definition of the premium subsidies, but also from the explanatory model from which these premium subsidies are derived.

There are some problems with this conventional approach. The implicit neglect of many explanatory variables tends to obscure important ethical and social implications. Moreover, if included C-variables and omitted R-variables are correlated, the conventional approach yields biased estimates of the former and hence biased calculations of the premium subsidies (Schokkaert and Van de Voorde, 2004). It has therefore been proposed to include all relevant variables in the explanatory model and to make only afterwards an explicit distinction between the C- and R-variables (Schokkaert *et al.*, 1998). The effects of the R-variables are then neutralized for the calculation of the premium subsidies, e.g. by setting them at an acceptable and identical level for all individuals[4]. We call this the "*explicit approach*" to risk adjustment. The comparison between the conventional and the explicit approach has until now mostly focused on the econometric aspects of estimating the risk adjustment formula. In this paper we go deeper into the different financial incen-

4. In Schokkaert *et al.* (1998) and Schokkaert and Van de Voorde (2004) the use of the mean is advocated - but there are obviously other possibilities.

tives for insurers created by both approaches. We will do this in the context of a very simple model with only two explanatory variables: one which can be seen as a C-variable, another which is interpreted as a R-variable.

Our analysis remains limited to describing the direct effects on financial incentives. We do not specify a full behavioral model of health insurers, leading to specific predictions of the market outcomes. The specification of a behavioral model that is relevant for the European situation is far from obvious. Most theoretical papers start from an assumption of profit-maximizing insurers and perfect consumer mobility. The first assumption is debatable with non-profit sickness funds and the second one is not at all corroborated in the empirical work with European data (Schut *et al.*, 2003). The literature (see, e.g., Encinosa, 1999; Sappington and Lewis, 1999; Glazer and Mc Guire, 2000, 2002; Jack, 2001) mainly focuses on adverse selection, while in the European systems with compulsory insurance and a well-defined benefits package, risk selection (i.e. cream-skimming) by the insurers is a more important problem[5]. To understand the latter phenomenon (and why it after all remains rather exceptional in countries like the Netherlands, Germany and Belgium), it is necessary to introduce reputation effects and to model carefully the selection instruments that are available to the insurers[6]. The formulation of a realistic and policy-relevant behavioral model is therefore well beyond the scope of this paper, but our analysis of "first-order" financial incentives is in any case a first and necessary step towards such a model. Since the limitations of our approach are most obvious for the second category of illegitimate variables, i.e. those directly linked to producer or insurer behavior, we will mainly focus on patient characteristics.

Although our model is a model of health insurance, our description of financial incentives is to some extent also relevant for the institutional setting of NHS-type systems. In the latter, risk adjustment and capitation have become important issues in the allocation of the financial means over different regions (Rice and Smith, 1999). Here also the distinction between legitimate and illegitimate risk adjusters has cropped up and has led to a discussion about the relative merits of the conventional and the explicit approaches (Carr-Hill *et al.*, 1994; Smith, Rice and Carr-Hill, 2001; Gravelle *et al.*, 2003). Risk selection (in our interpretation) and equality of access (in their interpretation) are closely linked.[7]

Section 2 introduces the model and considers the benchmark case in which the C- and R-variables are distributed independently in the population. However, the distinction between the conventional and the explicit approach only becomes relevant as soon as there is some correlation between C and R. In section 3 we analyze this more interesting situation for the case in which the R-variable is a patient characteristic that for equity reasons is not taken up in the risk adjustment formula. We argue that the traditional approach in that case creates distorted incentives for efficiency and equity and that the statistical arguments in favor of the explicit model can thus be supplemented by a more basic theoretical argument. In section 4 we focus on the more difficult and empirically more relevant case of manipulable morbidity information. Section 5 contains an empirical illustration of the problem with data for Belgium. Section 6 concludes.

---

5. Cream-skimming is the main focus in the papers by Ellis (1998), Marchand *et al.* (2003) and Barros (2003).
6. The interaction with supplementary health insurance schemes is one possibility (Kifmann, 2002), variations in the quality of non-health care services another (Marchand *et al.*, 2003).
7. Another possible application is the design of capitation schemes for the remuneration of providers. In that setting the danger of cream-skimming may even be larger.

2 A Simple Model

Let us assume that the expected expenditures of individual *i* can be written as a linear function of a compensation variable and a responsibility variable:

$$E_i = E_0 + \alpha C_i + \beta R_i \tag{1}$$

It would not be difficult to treat the variables C and R as continuous. In the real world, however, premium differentiation and risk selection are applied to discrete groups. We therefore assume that both C and R are discrete and, even more simply, binary variables. The C-variable is interpreted as related to morbidity. To focus ideas -and to make it easier to tell the story- we will interpret $C_i = 1$ if individual *i* is old and $C_i = 0$ if she is young. As we will give different interpretations to the R-variable in the following sections, we now remain general and specify $R_i = 1$ for "high responsibility" and $R_i = 0$ for "low responsibility". Given these interpretations, we impose $\alpha, \beta > 0$. Note that equation (1) does not include a random idiosyncratic term. Purely random elements (with mean 0) are irrelevant because we focus on expected expenditures. Other idiosyncratic factors can simply be seen as either C or (more probably) R variables and could be explicitly integrated into a model with more than two explanatory variables. Note also that the expenditure function (1) is additively separable. The complications when this is not the case are further discussed in Schokkaert *et al.* (1998). For the purpose of this paper we disregard these problems.

Given these assumptions, all insured can be classified into four mutually exclusive groups, denoted by the indices *cr* where $c \in (O, Y)$ (for old and young) and $r \in (H, L)$ for high and low responsibility respectively. We indicate the proportions of the different groups in the population by p_{cr}. Of course we have that $p_{YH} + p_{YL} + p_{OH} + p_{OL} = 1$. Moreover we introduce the obvious notation

$$\begin{aligned} p_O &= p_{OH} + p_{OL} \\ p_Y &= p_{YH} + p_{YL} \\ p_H &= p_{OH} + p_{YH} \\ p_L &= p_{OL} + p_{YL} \end{aligned} \tag{2}$$

Applying equation (1) and the definitions of C and R we can summarize expected expenditures for the four groups as follows:

$$\begin{aligned} E_{YL} &= E_0 \\ E_{YH} &= E_0 + \beta \\ E_{OL} &= E_0 + \alpha \\ E_{OH} &= E_0 + \alpha + \beta \end{aligned} \tag{3}$$

As described in the introduction, health insurers have to reimburse all health care expenditures of their members. Citizens pay premium contributions to that insurer and solidarity contributions to a general fund. The fund redistributes these solidarity contributions over the insurers by giving risk-adjusted premium subsidies. We denote the subsidy received for member *i* by N_i (which can be interpreted as "normative expenditures"). For simplicity we will assume that the sum of the premium subsidies (or normative expenditures) exactly covers total expenditures, i.e. $\sum_i N_i = \sum_i E_i$. In reality the total sum of premium subsidies is not sufficient to cover all expenses and the remainder has to be financed by the insurers through community-rated premiums. However, introducing such a fixed premium does not change any of our results and would complicate the notation.

For this simple model it is easy to compute the premium subsidies. We start from the basic idea that *R* is treated as an illegitimate risk adjuster. This means that insurers get different premium subsidies for the young and the old but that there is no differentiation for the high- and low responsibility-groups. Even then, however, in the *explicit* approach the full model (1) is estimated. To calculate the premium subsidies, the effect of the R-variable is neutralized by fixing it for everybody at its average value. This yields as premium subsidies

$$\text{(4)} \qquad \begin{aligned} \text{for the young: } N_Y^* &= E_0 + p_H\beta \\ \text{for the old: } N_O^* &= E_0 + \alpha + p_H\beta \end{aligned}$$

In the *conventional* approach the variable *R* is left out of the regression equation and its effect is partly taken up by the other coefficients. Alternatively (and equivalently) the premium subsidies are calculated as the cell means, where the cells are defined on the basis of the compensation variables only. This gives

$$\text{(5)} \qquad \begin{aligned} \text{for the young: } \tilde{N}_Y &= E_0 + \frac{p_{YH}}{p_Y}\beta \\ \text{for the old: } \tilde{N}_O &= E_0 + \alpha + \frac{p_{OH}}{p_O}\beta \end{aligned}$$

The point of this paper is to compare the consequences of these two options. Of course this problem becomes irrelevant if the C- and R-variables are distributed independently in the population, i.e. if $p_{cr} = p_c p_r \ \forall c, r$. In that case the premium subsidies in (4) and (5) are identical, i.e. the two approaches are fully equivalent. If *C* and *R* are not independently distributed, we can distinguish two cases. With some abuse of terminology (since we have assumed *C* and *R* to be discrete variables) we introduce the following definition:

DEFINITION 1. *The variables C and R are positively correlated if there are more high responsibility members among the aged, i.e. if* $\frac{p_{OH}}{p_O} > \frac{p_{YH}}{p_Y}$. *They are negatively correlated if there are less high responsibility members among the aged, i.e. if* $\frac{p_{OH}}{p_O} < \frac{p_{YH}}{p_Y}$.

In what follows we will give the results for both cases. To avoid unnecessary repetition, however, we will mainly focus on the case of positive correlation for the interpretation of our results. Throughout the paper, the $\sim$-symbol will be used to indicate the results for the conventional approach, while the *-superscript will refer to the explicit approach.

3 Patient Characteristics as Illegitimate Risk Adjusters

We will first analyze the case where the R-variable refers to a personal characteristic of the patient which is considered to be an illegitimate risk adjuster. Lifestyle variables are an obvious example and we will tell the story for smokers $(R_i = 1)$ versus non-smokers $(R_i = 0)$. This is only meant to be an example and we do not want to enter the discussion about the desirability of a differential treatment of smokers. In any case it seems realistic to suppose that "smoking" will not be introduced in any real-world risk adjustment system in the short run. An alternative example could have been the loyalty of the patient to his or her GP. It is well known (see Schokkaert and Van de Voorde, 2004, for some empirical evidence) that in systems where patients have a free choice of physician, those who remain loyal to one GP and who accept the gate-keeping role of that GP, have lower expenditures. What is crucial in these examples is that the R-variable, i.c. the smoking or medical shopping behavior of their members, cannot be directly controlled by the insurers.

3.1 Incentives for Risk Selection with Community Rated Premiums

As argued before, the whole point of risk adjustment is to avoid undesirable premium differentiation. However, before analysing the consequences of community rating, it is useful to consider first as a reference situation the hypothetical case in which insurers would be completely free to differentiate the premiums for different groups. It is then immediately clear that the actuarially fair premium Π for the different groups, *given* the system of premium subsidies, can be calculated as the difference between expected expenditures and premium subsidies[8]. This gives for the conventional model

8. We neglect transaction costs. Some of the calculated premiums will be negative. This unrealistic result follows from our assumption that the sum of premium subsidies exactly covers total health care expenses. As mentioned before, if this is not the case, one simply has to add a constant amount to all premiums. This would not change our conclusions.

(6)
$$\tilde{\Pi}_{YH} = \left(1 - \frac{p_{YH}}{p_Y}\right)\beta = \frac{p_{YL}}{p_Y}\beta$$
$$\tilde{\Pi}_{YL} = -\frac{p_{YH}}{p_Y}\beta$$
$$\tilde{\Pi}_{OH} = \left(1 - \frac{p_{OH}}{p_O}\right)\beta = \frac{p_{OL}}{p_O}\beta$$
$$\tilde{\Pi}_{OL} = -\frac{p_{OH}}{p_O}\beta$$

and for the explicit model

(7)
$$\Pi^*_{YH} = \Pi^*_{OH} = (1 - p_H)\beta$$
$$\Pi^*_{YL} = \Pi^*_{OL} = -p_H\beta$$

The interpretation of these results is easy. In both cases the premium is higher for smokers than for non-smokers - presumably a desirable result, given that the regulator assumes that people are responsible for their smoking behavior. In the explicit model the premium difference between smokers and non-smokers equals β, both for the old and for the young. This is the only premium difference. This is not true in the conventional model, however. In that model the premium difference between old smokers and old non-smokers is β, as is the difference between young smokers and young non-smokers. However, at the same time there is also a difference between young and old smokers, respectively non-smokers (with the direction of the difference depending on the direction of correlation between the C- and R-variables). This can be seen as an undesirable form of premium differentiation because it leads to differential treatment on the basis of a characteristic (age) for which people cannot be held responsible.

Let us now turn to the more realistic case of community rating. Call the community rated premium for an insurer *P*.[9] It is then obvious that insurers will make identifiable profits and losses on the different groups, given by

(8)
$$PRS_{cr} = P - \Pi_{cr} \quad \forall r, c$$

where *PRS* stands for "profits from risk selection". Combining (6), (7) and (8) we can immediately derive

Proposition 2. *In the explicit approach* $PRS^*_{OL} = PRS^*_{YL} > PRS^*_{OH} = PRS^*_{YH}$. *There are no incentives for differential treatment of the young and the old. However, in the conventional approach* $\widetilde{PRS}_{OL} > \widetilde{PRS}_{YL} > \widetilde{PRS}_{OH} > \widetilde{PRS}_{YH}$ *if there is a positive correlation between C and R, while* $\widetilde{PRS}_{YL} > \widetilde{PRS}_{OL} > \widetilde{PRS}_{YH} > \widetilde{PRS}_{OH}$ *if there is a negative correlation. Conventional risk adjustment does not remove the incentives for differential treatment of the young and the old.*

9. *P* will be zero if the risk profile for that insurer is the same as the risk profile of the population.

This proposition sketches a picture of the relative profitability of the different groups, i.e. of the financial incentives for differential treatment given to the insurers. The relative values of *PRS* will play an important role in any behavioral analysis (see, e.g., Glazer and McGuire, 2000, 2002). They are also closely related to the ideal of equality of access in NHS-type capitation systems. In the explicit approach there is a difference between smokers and non-smokers, but not between the old and the young. The result for the conventional approach is different. In the latter approach, the expenditure effect of smoking versus non-smoking is taken up in the premium subsidies for the young and the old. More specifically, in the case of positive correlation, i.e. if there are relatively more smokers among the old than among the young, the difference between the premium subsidies for the old and the young is increased. This means that both *within* the group of the smokers and *within* the group of the non-smokers the old, i.e. the group that smokes most on average, are relatively more attractive.

We can reasonably speculate about the behavioral consequences of these differences. With the explicit approach financial losses are made on smokers, gains on non-smokers. If the relative gains are small, insurers will not react to the difference. If the relative gains and losses are considerable, however, insurers have two options. Since they cannot directly influence the smoking behavior of their members, they most probably will try to attract non-smokers rather than smokers, i.e. they will engage in cream-skimming. Or, they may set up a general campaign to discourage smoking. In any case they do not have any incentive to discriminate between the young and the old. In the conventional model, however, there are differential gains and losses associated with the C-variable. More specifically, in the case of positive correlation insurers get incentives to focus their cream-skimming efforts on the old (rather than the young) non-smokers and to explicitly deter the young (rather than the old) smokers. This implies that there will be risk selection on the basis of a variable which is beyond the control of the citizens. Such differential treatment is socially undesirable. This offers a clear argument in favor of the explicit model.

There remains an important *caveat* to be made, however. Until now we implicitly assumed that insurers can differentiate their treatment of patients along both the C- and the R-dimension at the same time - and that without risk adjustment it would be profitable for them to do so. A minimal requirement for this is that they can observe the R-variable (in our example smoking behavior) at the individual level. This is far from evident. Let us therefore now assume that it is not possible for the insurers to differentiate on the basis of the R-variable and that they are restricted to base their policy on the C-variable, i.e. in this case on age[10]. The profits from risk selection will then depend on the difference between the expected expenditures and the premium subsidies for the old and the young respectively. These expected expenditures will be a weighted average of the expenditures of smokers and non-smokers, with the weights depending on the fractions of young and old (non)smokers in the membership of the insurer. For simplicity we assume that these are equal to the population proportions in (2)[11].

10. The alternative extreme assumption -that insurers only select on the basis of R-variables and do not select on the basis of the C-variable at all- is not very interesting. Since this kind of cream-skimming is not problematic from an ethical point of view, there is no real need for intervention by the regulator.
11. Weakening this assumption would not change our conclusions as long as the sign of the correlation as defined in definition 1 remains the same for the insurer and for the population.

In the conventional model premium subsidies (5) for the young and the old are exactly equal to expected expenditures by definition, i.e.

$$\widetilde{PRS}_Y = \widetilde{PRS}_O = 0 \tag{9}$$

where the subscripts indicate that differential treatment is only feasible on the basis of the C-variable. For the explicit model, however, we have to subtract expected expenditures from the premium subsidies given by eq. (4). This yields

$$\begin{aligned} PRS_Y^* &= \beta\left(p_H - \frac{p_{YH}}{p_Y} \right) \\ PRS_O^* &= \beta\left(p_H - \frac{p_{OH}}{p_O} \right) \end{aligned} \tag{10}$$

We therefore get

Proposition 3. *If insurers can only differentiate (or select risks) on the basis of C (and not of R), then the conventional model removes all incentives for risk selection, while in the explicit model* $PRS_Y^{*C} > PRS_O^{*C}$ *if there is a positive correlation between C and R and* $PRS_Y^{*C} < PRS_O^{*C}$ *if this correlation is negative.*

This result reflects the important insight that the information taken up in the risk adjustment system should mimic as accurately as possible the information that is available to the insurers. Note its strong implications. Statistical considerations do not offer the decisive argument concerning the choice of variables to include in the risk adjustment system. In fact, Proposition 3 implies that it may be preferable to set up a risk adjustment system with biased coefficient estimates even if the regulator disposes of the necessary information to remove the omitted variables bias.

At the same time, however, the assumption that insurers cannot differentiate at all on the basis of the R-variable is a rather extreme one. To keep to our example: smoking behavior is already used to set premiums by private lifetime insurance companies. And even if insurers do not have at their disposal information on individual smoking behavior, they still may launch general campaigns towards smokers and non-smokers and try to make themselves more attractive for one of both groups. In a more complete model intermediate assumptions, e.g. that risk selection is easier on the basis of C than on the basis of R, should also be considered. A satisfactory analysis of the trade-offs involved would require the elaboration of a complete behavioral model and a fully specified social welfare function. Note that in the choice between differentiating on the basis of C or R, reputation costs should not be neglected in countries with a long tradition in compulsory health insurance and where the population attaches much importance to equity considerations. This suggests that the danger of introducing "too many" explanatory variables in the explicit model should perhaps not be exaggerated. We will return to the questions about the relative incentives for risk selection against different groups in section 4 on manipulable morbidity information.

3.2 **Premium Differentiation on the Basis of the R-Variable**

As discussed before, the main rationale for community rating with risk adjustment is to guarantee equal treatment for all. However, as soon as we introduce a distinction between C- and R-variables and hold people responsible for differences in the latter, there is no longer any compelling ethical reason for forbidding differential treatment and premium differentiation on the basis of the R-variable. If the R-variable is taken out of the risk adjustment system, while it is not directly under the control of the insurers, it is reasonable to allow for the opportunity of differentiating the premiums for different R-groups (here smokers and non-smokers), while at the same time still prohibiting premium differentiation for different C-groups (here the young and the old). This possibility is also suggested by, e.g., van de Ven and Ellis (2000).

Let us call the resulting premiums Π_H^R and Π_L^R for smokers and non-smokers respectively (where the superscript indicates that only premium differentiation on the basis of the R-variable is possible). If we accept as before that the premiums for smokers and non-smokers are calculated as the difference between their expected expenditures and the average premium subsidies received, Π_H^R and Π_L^R will be a weighted average of the expressions in (6) and (7). Using again the assumption that the weights can be based on the population proportions, we get (for $r = H, L$)

$$\Pi_r^R = \frac{p_{Or}}{p_r}\Pi_{Or} + \frac{p_{Yr}}{p_r}\Pi_{Yr} \qquad (11)$$

The remaining incentives for risk selection can then be written as

$$PRS_{cr}^R = \Pi_r^R - \Pi_{cr} \quad \forall r, c \qquad (12)$$

The restriction on premium differentiation imposed by the regulator does not change anything in the explicit approach, since (7) and proposition 2 show that insurers will not discriminate between the old and the young, even if they are allowed to do so. We therefore get

$$\Pi_H^{R*} = \Pi_{YH}^* = \Pi_{OH}^* > \Pi_L^{R*} = \Pi_{YL}^* = \Pi_{OL}^* \qquad (13)$$

and, more specifically,

$$\Pi_H^{R*} - \Pi_L^{R*} = \beta \qquad (14)$$

The difference between the premiums for smokers and non-smokers is exactly equal to the differences in expected expenditures for both groups. It immediately follows that

$$PRS_{YH}^{R*} = PRS_{OH}^{R*} = PRS_{YL}^{R*} = PRS_{OL}^{R*} = 0 \qquad (15)$$

After allowing for premium differentiation between the smokers and the non-smokers there are no incentives for risk selection left.

Things are very different in the conventional approach, however. Using (11), we derive immediately

$$\tag{16} \begin{aligned} \tilde{\Pi}_H^R &= \frac{\beta}{p_H}\left(p_{OH}\frac{p_{OL}}{p_O} + p_{YH}\frac{p_{YL}}{p_Y}\right) \\ \tilde{\Pi}_L^R &= -\frac{\beta}{p_L}\left(p_{OL}\frac{p_{OH}}{p_O} + p_{YL}\frac{p_{YH}}{p_Y}\right) \end{aligned}$$

and, using (12),

$$\tag{17} \begin{aligned} \widetilde{PRS}_{YH}^R &= -\beta\frac{p_{OH}}{p_H}\left(\frac{p_{OH}}{p_O} - \frac{p_{YH}}{p_Y}\right) \\ \widetilde{PRS}_{YL}^R &= -\beta\frac{p_{OL}}{p_L}\left(\frac{p_{OH}}{p_O} - \frac{p_{YH}}{p_Y}\right) \\ \widetilde{PRS}_{OH}^R &= \beta\frac{p_{YH}}{p_H}\left(\frac{p_{OH}}{p_O} - \frac{p_{YH}}{p_Y}\right) \\ \widetilde{PRS}_{OL}^R &= \beta\frac{p_{YL}}{p_L}\left(\frac{p_{OH}}{p_O} - \frac{p_{YH}}{p_Y}\right) \end{aligned}$$

The information in the equations (13), (15), (16) and (17) can now be used to show

PROPOSITION 4. *If premium differentiation on the basis of the R-variable is possible (while it remains forbidden on the basis of the C-variable), both approaches lead to higher premiums for high-responsibility individuals. In the explicit approach there are no incentives for risk selection left. However, in the conventional approach* $\widetilde{PRS}_{OL}^R > \widetilde{PRS}_{OH}^R > \widetilde{PRS}_{YL}^R > \widetilde{PRS}_{YH}^R$ *if there is a positive correlation between C and R, while* $\widetilde{PRS}_{YL}^R > \widetilde{PRS}_{YH}^R > \widetilde{PRS}_{OL}^R > \widetilde{PRS}_{OH}^R$ *if there is a negative correlation.*

To interpret the results for the conventional approach, let us focus on the case of positive correlation and compare the rankings of PRS in propositions 2 and 4, i.e. in the case without and with premium differentiation between smokers and non-smokers respectively. The most striking change is that it is now unambigously preferable to focus on the old rather than on the young, and independent of the fact whether they are smoking or not. This follows from the changes in the ranking of $\widetilde{PRS}_{OH}^R$ and $\widetilde{PRS}_{YL}^R$. Without premium differentiation young non-smokers are more profitable for the insurers than old smokers, because the difference in the premium subsidies for the young and the old is a weighted average and therefore not sufficient to compensate for the differences in expected expenditures between these two extreme groups. With premium differentiation things change because the lower premiums for non-smokers overcompensate the difference in expected expenditures.

In our view, proposition 4 suggests the most adequate policy in this case of patient characteristics as illegitimate risk adjusters. We endorse the view that there is no reason not to give the insurers the option to differentiate their premiums on the basis of an illegitimate characteristic. That option, however, will only remove all the incentives for risk selection *provided that the explicit approach is used to calculate the premium subsidies*, i.e. that the information on these characteristics is introduced in the equations explaining expected expenditures and neutralized *afterwards* for the calculation of the premium subsidies.

4 Manipulable Morbidity Information as an Illegitimate Risk Adjuster

Let us now turn to the treatment in the risk adjustment formula of variables which are an indicator of morbidity but can at the same time be manipulated or influenced by the insurers. While insurers observe past health care expenditures or diagnoses and can exploit this information for risk selection purposes, introducing this information in the risk adjustment system would create incentives for manipulation and dilute the incentives for efficiency. In most cases regulators consider the latter danger as sufficiently important to treat these variables as illegitimate risk adjusters and to leave them out of the definition of normative expenditures[12]. This then again raises the question of the choice between the conventional and the explicit model. Taking for granted that these variables are neutralized for the calculation of the premium subsidies, should they then also be left out from the explanatory model? This "conventional" approach is certainly dominant in current practice.

Formally, the problem is completely analogous to the one analysed in the previous section. The R-variable now is a morbidity related variable, however. We will talk about "high severity" and "low severity" respectively. This changes the interpretation of the problem considerably. More specifically, it is hard to imagine that the insurers would get the option to differentiate the premiums for "high severity" and "low severity" members, because individual enrollees cannot be held ethically responsible for the degree of severity of their illness. The solution described in section 3.2 is therefore not relevant in this setting. The whole focus of the exercise will be on minimizing the incentives for risk selection. Since all examples of differential treatment are now equally undesirable, the regulator's objective can be formulated as minimizing the dispersion in the *PRS* for the different groups. We take the difference between *PRS* for the most and the least profitable group as a relevant measure for this dispersion[13].

12. Marchand *et al.* (2003) summarize the arguments for the important example of past expenditures.
13. The choice of this indicator is important for the results. An alternative criterion (e.g. the variance of PRS) would not necessarily lead to the same conclusions. While we feel that our focus on the largest absolute difference is defensible, more research is needed to explore the consequences of such alternative criteria. Again, a full analysis would require the specification of a complete behavioral model.

At first sight, a simple intuition in favor of the conventional model could then be the following. If we leave out the severity variable from the premium subsidies, we know that we cannot avoid the problem of cream skimming. However, one could argue that, by leaving the severity variable out of the explanatory model also, the cream-skimming problem is mitigated because at least part of the severity effect will be taken over by the coefficients of the variables which are included, i.e. in our simple example by the age variable.[14] This simple intuition is exactly captured by Proposition 3. The conventional model removes all incentives for risk selection on the basis of age, while the explicit model makes selection on the basis of age profitable because of the neglected correlation between age and severity. Combining (9) and (10) we get

$$(18) \qquad PRS_Y^* - PRS_O^* = \beta\left(\frac{p_{OH}}{p_O} - \frac{p_{YH}}{p_Y}\right) > \widetilde{PRS}_Y - \widetilde{PRS}_O = 0$$

where the subscripts indicate that risk selection takes place on the basis of the C-variable only.

The analogy with Proposition 3, however, immediately suggests that this simple intuition in favor of the conventional model is wrong or at least one-sided. Why would insurers limit themselves to risk selection on the basis of age if they have the severity information readily available? It seems much more natural to assume that they will exploit this information. This brings us directly into the setting of section 3.1 and more specifically Proposition 2. The latter proposition confirms that cream-skimming incentives are unavoidable if one treats morbidity variables as illegitimate risk adjusters. Indeed, the profits from risk selection are not equalized for the different groups in the explicit approach nor in the conventional approach. Following then our strategy of looking at the dispersion of *PRS* between the most and the least profitable group, we get for the conventional model in the case of positive correlation:

$$(19) \qquad \widetilde{PRS}_{OL} - \widetilde{PRS}_{YH} = \beta\left(\frac{p_{YL}}{p_Y} + \frac{p_{OH}}{p_O}\right) = \beta\left(1 - \frac{p_{YH}}{p_Y} + \frac{p_{OH}}{p_O}\right) > \beta$$

while for the explicit model

$$(20) \qquad PRS_{OL}^* - PRS_{YH}^* = \beta$$

We can therefore conclude

PROPOSITION 5. *If one decides to leave out manipulable morbidity variables from the risk adjustment formula the resulting range for profitable cream-skimming is larger in the conventional than in the explicit model.*

14. This reasoning is very similar to the one which was used originally by the York-group (Carr-Hill *et al.*, 1994) to advocate leaving out medical supply variables from the estimated equations.

We think that proposition 5 conveys an important insight. It suggests that the explicit approach is to be preferred if one wants to minimize the incentives for risk selection, i.e. that it is advisable to introduce variables such as past health care expenditures and manipulable diagnostic information into the regression equations and neutralize their effects afterwards rather than leaving them out altogether. This conclusion goes directly against current practice in the whole risk adjustment literature. We will show in the next section that our point may be relevant in real-world applications.

5 An Empirical Illustration

In the Belgian health insurance system the calculation of the risk adjusted premium subsidies is based on a regression equation estimated with individual data[15]. Until now no direct diagnostic information is available, apart from some categories of disability. However there are some variables which are closely linked to morbidity. One of these is the classification of members of sickness funds as "being chronically ill".[16] While average health care expenditures of those classified as "chronically ill" are much larger than those of the overall population, sickness funds can exert some influence on this classification. Including the indicator of "chronical illness" gives the sickness funds a strong incentive to reclassify some of their members and immediately confronts us with the questions raised in the previous section.

Of course in actual reality the regression equation and the risk adjustment formula contain more than two variables. We did not attempt to treat the full problem with *n* variables in this paper because the large number of correlations between the variables makes an easy interpretation impossible. This paper can be seen as a partial approach under the implicit assumption that some correlations strongly dominate the other. The complete Belgian model is described in the Appendix. This Appendix shows that our two variables-model is indeed a useful, although very simple, approximation of reality. We will therefore stick to this simple approach in this section. Compared to the full results given in the Appendix, this can be interpreted as a *ceteris paribus* approach: we look at the effects of two variables conditional on the values of all the others. The first (C) variable is age, with $C_i = 1$ if individual i is at least 70 years old, while $C_i = 0$ if she is younger. The second variable takes the value 1 if the individual is classified as "being chronically ill" and takes the value 0 otherwise. Since "being chronically ill" can to some extent be manipulated by the sickness funds, we consider it as a R-variable. We use again the self-explanatory subscripts Y(oung), O(ld), H(igh severity) and L(ow severity). The relevant proportions of the different groups in the Belgian population

15. The main features of the risk adjustment system are described in the Appendix. More detailed information on the Belgian health insurance system can be found in Schokkaert and Van de Voorde (2003).
16. For a member to be classified as "chronically ill", two conditions have to be satisfied. First, her personal copayments must have reached a given threshold level over two consecutive years. Second, she must be classified as being in need of care (e.g. home care or special physiotherapy - see Appendix).

(excluding the self-employed) are given in Table 1. It shows that our empirical illustration is an example of positive correlation between C and R. All estimations are based on a representative sample of 438,401 individuals. Table 2 shows the estimation results for the explicit and the conventional approach. In the latter the R-variable is left out of the equation. Normative expenditures (or premium subsidies) are given in the fifth and the sixth row of Table 3. For the explicit approach they follow from application of eq. (4) to the estimates in the left hand column of Table 2. For the conventional model they follow directly from the estimates in the right hand column of Table 2. They could also be derived from the estimates in the left-hand column through application of eq. (5). It is clear that both approaches give a significant difference: the premium subsidies for the old are in the conventional model almost 230 Euro larger than in the explicit model. The other figures in Table 3 give the profits from risk selection in the different cases described in the previous section. It is immediately obvious that the treatment of "being chronically ill" as an illegitimate risk adjuster and leaving it out from the calculation of normative expenditures creates huge financial incentives for risk selection in both approaches: this is not surprising given the large value for β in Table 2. In fact, this result completely swamps all the other effects. Note that this large value for β is really a two-edged sword. On the one hand, it points to a potentially important equity problem. On the other hand, it suggests that the problem of misclassification and the resulting cost increases and efficiency losses may be very important. The trade-off for the policy maker is a difficult one.

At the same time all the theoretical results from the previous sections are illustrated by the figures in Table 3. More specifically, if insurers concentrate only on the age dimension, the conventional approach might be preferable. However, this assumption is highly unrealistic in this setting. Given that the insurers themselves are involved in the classification of members as "being chronically ill", they have ample opportunities to exploit this information. Reputation effects may even be less detrimental than with age discrimination, because the classification process

TABLE 1
Proportions in the population

p_O	p_{OH}	p_{OL}	p_Y	p_{YH}	p_{YL}	p_H	p_L
12.67%	0.60%	12.07%	87.33%	0.54%	86.80%	1.13%	98.87%
$p_{OH}/p_O = 0.04706$				$p_{YH}/p_Y = 0.00613$			

TABLE 2
Estimation results for the explicit and the conventional approach

	Explicit model	Conventional model
E_0	683 (5.40)	723 (5.50)
α (age)	2185 (15.27)	2448 (15.45)
β (severity)	6421 (48.03)	
R^2	.091	.054

of "being chronically ill" is subtle and hardly observed by the general public. It is therefore interesting to note that the explicit model removes all the incentives to discriminate between the young and the old within the high and the low severity group. More importantly, the dispersion in the profits from risk selection between the most profitable (the old not in need of care) and the least profitable (the young in need of care) groups is smaller in the explicit model (6421 Euro) than in the conventional model (6684 Euro): the difference of more than 260 Euro between the two is certainly not negligible.

TABLE 3

Premium subsidies and incentives for risk selection (in Euro)

	Explicit model	Conventional model
E_{YL}	683	
E_{YH}	7104	
E_{OL}	2868	
E_{OH}	12158	
N_Y	756	723
N_O	2941	3171
PRS_{YH}	-6348	-6381
PRS_{YL}	73	39
PRS_{OH}	-6348	-6119
PRS_{OL}	73	302
PRS_Y	33	0
PRS_O	-230	0
$PRS_Y - PRS_O$	263	0
$PRS_{OL} - PRS_{YH}$	6421	6684

6 Conclusion

The equity results of managed competition in health insurance depend crucially on the quality of the risk adjustment system. The regulator first has to take a decision on what variables to include in the system, the so-called legitimate risk-adjusters. This is mainly an ethical and/or political choice. In a second stage, he then has to decide about the weights to attach to the included variables. In practice there does not seem to be an alternative to the procedure of deriving these weights from the observation of actual expenditure patterns. However, actual health care expenditures are codetermined by variables which are not included in the risk adjustment system for equity or efficiency reasons, the so-called illegitimate risk adjusters. How then to go from observed to normative expenditures?

Current practice is to neglect these illegitimate risk adjusters, more specifically to estimate a regression model with only legitimate risk adjusters included and then compute normative expenditures or risk-adjusted premium subsidies as the predicted expenditures generated by this regression. It has been argued before that this conventional approach leads to biased estimates of the effects of the included variables (Gravelle *et al.*, 2003; Schokkaert and Van de Voorde, 2004). In this paper we presented some additional theoretical arguments in favor of the alternative so-called explicit approach, in which all empirically relevant variables are included in the regression equation but the effects of the illegitimate risk adjusters are neutralized afterwards by putting them at their mean value.

We first looked at the case of patient characteristics for which individuals should be held ethically responsible. Lifestyle variables could be an example. We show that the conventional approach does not remove all incentives for undesirable risk selection on the basis of the legitimate risk adjusters, e.g. age. In the explicit approach, however, there are no incentives for differential treatment of the young and the old. This becomes especially relevant when we consider the policy choice of allowing premium differentiation on the basis of the illegitimate risk adjusters. This policy only works in the explicit approach, where all incentives for risk selection are removed. In the conventional approach, however, there remain incentives for undesirable risk selection between the young and the old. We then considered the important case of morbidity-related variables which are left out for the calculation of the premium subsidies because they can be influenced by the insurers. Past health care expenditures and manipulable diagnostic information are obvious examples. Leaving out this information necessarily creates incentives for risk selection. However, we have shown that the resulting range for profitable cream-skimming is larger in the conventional than in the explicit model. The empirical and social relevancy of the choice between both has been illustrated with real-world Belgian data.

It is useful to compare our conclusions with some of the recent theoretical literature on "optimal" risk adjustment (Glazer and McGuire, 2000). Although we remain close to the practice of basing the risk adjustment formula on the estimated coefficients from a regression model, we share with them the idea that it is necessary to focus on incentives rather than on the statistical performance of the model. Moreover, we also derive the same conclusion that it is not optimal to put the premium subsidies for different patient groups equal to their average expenditures. They derive from a fully specified behavioral model with adverse selection the conclusion that optimal risk adjustment should pay higher than conventional risk adjustment for persons with the "bad" signal (the old), and lower than conventional risk adjustment for persons with the "good" signal (the young). Although their reasoning is different, this conclusion bears a striking similarity with the interpretation of our Proposition 3 that it may be advisable to introduce biased coefficient estimates in the risk adjustment model. Exploring further the links between "optimal" risk adjustment and what we have called the "explicit" approach is an interesting topic for further research.

Our analysis was based on a simple model with two binary variables. Since our results basically reflect the consequences of omitted variables bias, going from binary to continuous variables is easy and would not change any of our results. Moving to a setting with n variables would complicate the analysis and render interesting interpretations difficult. We think that our simple partial approach captures the main insights. The crucial weakness of our paper, however, is the lack of a complete model of insurer behavior. We basically limit ourselves to the preliminary step of describing first-order financial incentives. Future work should certainly focus on a more elaborated behavioral analysis. This requires careful thinking about the working of health insurance markets, including reputation effects, and about the objective function of non-profit sickness funds.

All in all, we think that the case in favor of the explicit approach is pretty strong. Note that this implies a breach with the current practice. At the same time, however, it does not raise many additional practical complications. It simply puts the empirical analyst in the position where she should be: that of trying to find the best possible explanatory model of health care expenditures while relegating the normative choices to the regulator in a second stage.

References

BARROS P. (2003). – "Cream-Skimming, Incentives for Efficiency and Payment System", *Journal of Health Economics*, 22: pp. 419-443.

CARR-HILL R., SHELDON T., SMITH P., MARTIN S., PEACOCK S. and HARDMAN G. (1994). – "Allocating Resources to Health Authorities: Development of Method for Small Area Analysis of Use of Inpatient Services", *British Medical Journal*, 309: pp. 1046-1049.

DWORKIN R. (1981). – "What is Equality? Part 2: Equality of Resources", *Philosophy and Public Affairs*, 10: pp. 283-345.

ELLIS R. (1998). – "Creaming, Skimping and Dumping: Provider Competition on the Intensive and Extensive Margins", *Journal of Health Economics*, 17: pp. 537-555.

ENCINOSA W. (1999). – "The Economic Theory of Risk Adjusting Capitation Rates." US Department for Health Care Policy and Research, mimeo.

FLEURBAEY M. and MANIQUET F. (2003). – "Compensation and Responsibility", forthcoming in *Handbook of Social Choice and Welfare*.

GLAZER J. and MCGUIRE T. (2000). – "Optimal Risk Adjustment in Markets with Adverse Selection: an Application to Managed Care", *American Economic Review*, 90: pp. 1055-1071.

GLAZER J. and MCGUIRE T. (2002). – "Setting Health Plan Premiums to Ensure Efficient Quality in Health Care: Minimum Variance Optimal Risk Adjustment", *Journal of Public Economics,* 84: pp. 153-173.

GRAVELLE H., SUTTON M., MORRIS S., WINDMEIJER F., LEYLAND A., DIBBEN C. and MUIRHEAD M. (2003). – "Modelling Supply and Demand Influences on the Use of Health Care: Implications for Deriving a Needs-Based Capitation Formula", *Health Economics,* 12: pp. 985-1004.

JACK W. (2001). – "Controlling Selection Incentives when Health Insurance Contracts are Endogenous", *Journal of Public Economics*, 80:25-48.

KIFMANN M. (2002). – "Community Rating in Health Insurance and Different Benefit Packages", *Journal of Health Economics*, 21: pp. 719-737.

MARCHAND M., SATO M. and SCHOKKAERT E. (2003). – "Prior Health Expenditures and Risk Sharing with Insurers Competing on Quality", *Rand Journal of Economics,* 34: pp. 647-669.

RICE N. and SMITH P. (1999). – "Approaches to Capitation and Risk Adjustment in Health Care: an International Survey", Centre for Health Economics, University of York.

SAPPINGTON D. and LEWIS T. (1999). – "Using Subjective Risk Adjusting to Prevent Patient Dumping in the Health Care Industry", *Journal of Economics and Management Strategy*, 8: pp. 351-382.

Schokkaert E., Dhaene G. and Van de Voorde C. (1998). – “Risk Adjustment and the Trade-Off Between Efficiency and Risk Selection: an Application of the Theory of Fair Compensation”, *Health Economics*, 7: pp. 465-480.

Schokkaert E. and Van de Voorde C. (2003). – “Belgium: Risk Adjustment and Financial Responsibility in a Centralised System”, *Health Policy*, 65: pp. 5-19.

Schokkaert E. and Van de Voorde C. (2004). – “Risk Selection and the Specification of the Conventional Risk Adjustment Formula”, *Journal of Health Economics,* 23: pp; 1237-1259.

Schut F., Gress S. and Wasem J. (2003). – “Consumer Price Sensitivity and Social Health Insurer Choice in Germany and the Netherlands”, *International Journal of Health Care Finance and Economics*, 3: pp. 117-138.

Smith P., Rice N. and Carr-Hill R. (2001). – “Capitation Funding in the Public Sector”, *Journal of the Royal Statistical Society A*, 164: pp. 217-257.

van de Ven W., Beck K., Buchner F., Chernichovsky D., Gardiol L., Holly A., Lamers L., Schokkaert E., Shmueli A., Spycher S., Van de Voorde C., van Vliet R., Wasem J. and Zmora I. (2003). – “Risk Adjustment and Risk Selection on the Sickness Fund Insurance Market in Five European Countries”, *Health Policy,* 65: pp. 75-98.

van de Ven, W. and Ellis R. (2000). – “Risk Adjustment in Competitive Health Plan Markets. In: Culyer, A., Newhouse, J. (eds.)”, *Handbook of Health Economics 1,* North-Holland, Amsterdam, pp. 755-845.

Appendix: The Belgian Risk Adjustment System

Belgium has a system of compulsory health insurance, covering the entire population and with a very broad benefits package. The management and administration of insurance is left to five non-governmental non-profit organisations -the sickness funds- and one public fund. The five national associations group a total of about a hundred local sickness funds. While membership of a sickness fund is compulsory, every individual can enrol in the sickness fund of her choice. The overall structure is very close to the model of managed competition, as described in Figure 1 in the text. The premium contributions are very low, however.

Before 1995 the central fund reimbursed all health care expenditures to the sickness funds. In 1995 this system was replaced by a mixed reimbursement formula for distributing the fixed health care budget over the sickness funds. The financing system is a weighted average of normative (risk-adjusted) and actual expenditures. Since the regulator wanted to introduce the financial responsibility of the sickness funds in a slow and cautious way, the weight given to the prospective part was set at a very low value (0.10) in the beginning. At this moment this value equals 0.30.

The focus on equity and equal access in Belgium has led to a complicated definition of normative expenditures with a long list of risk adjusters. Given this long list, including some continuous variables, the weights of the different explanatory variables were derived from a regression analysis. The first column of table A presents the estimation results for the general regime (i.e. excluding the self-employed) which is used to close the accounts of 2002 and 2003. The data are administrative data from the sickness funds and refer to the year 1998. A 5% random sample was drawn from the total population of the general regime (n=438.401). Except for some environmental variables, all explanatory variables are binary. The dependent variable equals total medical expenditures without expenditures for ambulatory medicines. Its mean is 1033 Euro.

The explanatory variables may be broadly classified into three groups. The first includes demographic and socio-economic variables such as age and sex, social group, isolation and specific categories of people who benefit from a social exemption of the copayments. The second group of explanatory variables provides direct or indirect information on the individual health status: mortality, chronically ill, disablement and invalidity. A last group contains environmental variables such as urbanisation and medical supply. The latter variables are based on aggregate data at the level of the municipality.

The results in column 1 are in line with other studies of health care expenditures. We find the classical U-shaped relationship with age for both men and women meaning that the effect of age on medical expenditures is higher for the very young and for the elderly. The coefficients for the variables relating to the social status of the insured indicate that being a widow or an orphan, living in isolation and belonging to certain groups with preferential reimbursement are associated with higher medical costs. The variables containing information on the individual health status have a major influence on health care costs: people who died, were disabled or invalid or chronically ill during the year and to a lesser extent people living in isolation have substantially higher expenditures. Also the regional factors

-the indicators for urbanisation and medical supply, measured through the regional density of practitioners- turn out to have a significant positive effect on medical expenditures. Since in an explicit approach a distinction has to be made between C- and R-variables, the Belgian regulator decided that the sickness funds should be held responsible for differences in medical expenditures that are caused by differences in medical supply. Hence medical supply is the only variable taken out of the formula to calculate the risk-adjusted premium subsidies.

It is a point of discussion whether the indicators for "being chronically ill" will still be used to calculate the premium subsidies in the future. It has been argued that they can be influenced to some extent by the sickness funds. The second column presents the estimation results of the complete Belgian model but with the manipulable morbidity variables left out of the estimation. These are the results which would result from following what we called the conventional approach. To illustrate that the two-variables model in the paper captures the main issue, we concentrate on the differences in the age effects between the conventional and explicit approach. Figures 2 and 3 show the difference between the conventional and the explicit approach for the elderly: when the manipulable morbidity variables for the chronically ill are left out of the equation, the age effects for the elderly, both for men and women, capture part of their effect.

FIGURE 2
Age effects for women

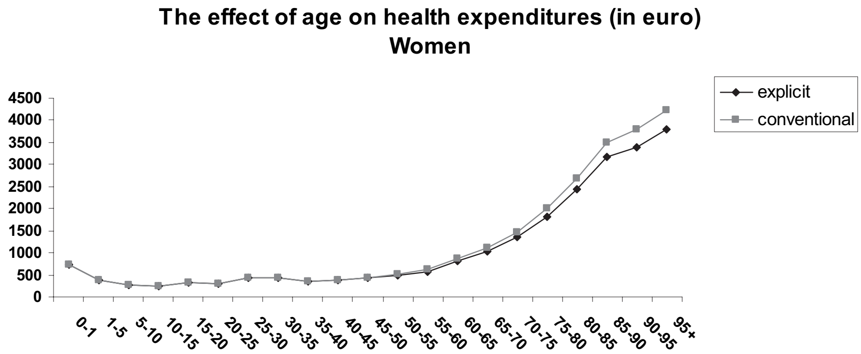


FIGURE 3
Age effects for men

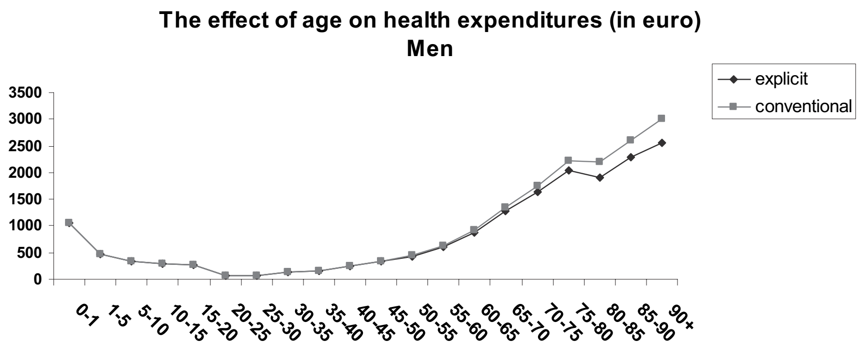


TABLE A
Estimation results for the explicit and the conventional approach (in Euro)

	Explicit		Conventional	
man of age 0-1	1.059	(62)	1.062	(63)
man of age 1-5	463	(31)	474	(31)
man of age 5-10	326	(27)	342	(27)
man of age 10-15	287	(27)	294	(28)
man of age 15-20	262	(27)	268	(28)
man of age 20-25	60	(28)	69	(28)
man of age 25-30	58	(27)	62	(27)
man of age 30-35	134	(25)	142	(26)
man of age 35-40	167	(25)	167	(25)
man of age 40-45	252	(26)	257	(26)
man of age 45-50	337	(26)	346	(27)
man of age 50-55	433	(28)	439	(28)
man of age 55-60	615	(31)	638	(32)
man of age 60-65	886	(30)	916	(31)
man of age 65-70	1.289	(30)	1.348	(31)
man of age 70-75	1.636	(34)	1.744	(34)
man of age 75-80	2.031	(39)	2.217	(39)
man of age 80-85	1.902	(60)	2.194	(61)
man of age 85-90	2.288	(76)	2.608	(77)
man of age >90	2.550	(128)	3.017	(130)
woman of age 0-1	733	(61)	733	(63)
woman of age 1-5	371	(31)	382	(32)
woman of age 5-10	264	(28)	277	(28)
woman of age 10-15	244	(28)	253	(29)
woman of age 15-20	320	(28)	324	(28)
woman of age 20-25	293	(28)	303	(28)
woman of age 25-30	423	(26)	428	(27)
woman of age 30-35	437	(25)	442	(25)
woman of age 35-40	351	(25)	357	(25)
woman of age 40-45	372	(25)	378	(26)
woman of age 45-50	430	(26)	445	(27)
woman of age 50-55	477	(27)	509	(28)
woman of age 55-60	577	(30)	616	(31)
woman of age 60-65	814	(29)	868	(30)
woman of age 65-70	1.042	(29)	1.113	(29)
woman of age 70-75	1.351	(31)	1.463	(31)
woman of age 75-80	1.822	(34)	2.005	(34)
woman of age 80-85	2.430	(46)	2.695	(47)
woman of age 85-90	3.168	(51)	3.508	(51)
woman of age 90-95	3.382	(75)	3.784	(77)

woman of age >95	3.784	(149)	4.224	(151)
widow/orphan	239	(27)	260	(28)
preferential reimbursement	503	(19)	493	(19)
mortality	6.053	(50)	6.182	(50)
disability: 1 year	1.549	(25)	1.553	(25)
isolation	223	(13)	209	(13)
urbanisation	15	(4)	12	(4)
medical supply	22	(5)	22	(5)
Social exemption				
-allowance for handicapped	1.321	(51)	1.478	(52)
-increased child allowance	2.891	(95)	3.451	(96)
-subsistence allowance	283	(39)	400	(38)
Chronically ill				
-entitled to home care (forfait B)	8.211	(101)		
-entitled to home care (forfait C)	10.073	(174)		
-special physiotherapy	3.543	(59)		
-allowance for third-party assistance	1.654	(282)		
-integration subsidy for handicapped	886	(79)	1.573	(81)
-increased allowance for disabled and invalids	1.029	(225)		
-allowance for the elderly	2.987	(90)	4.078	(91)
-allowance for third-party assistance (for handicapped)	1.257	(121)		
Diagnoses invalids				
-infectious and parasitary diseases	1.654	(315)	1.833	(321)
-tumours	2.314	(146)	2.324	(149)
-endocrine, nutritional and metabolic diseases and immunity disorders	2.559	(212)	2.363	(216)
-blood diseases and diseases of the hematopoietic organs	14.387	(842)	14.068	(858)
-psychological disorders	2.488	(76)	2.262	(77)
-respiratory diseases	648	(157)	731	(160)
-diseases of the digestive system	1.898	(174)	1.724	(177)
-urogenital diseases	9.531	(333)	9.257	(339)
-diseases of the motory system and of the connective tissue	-548	(78)	-669	(80)
-congenital defects	1.420	(379)	1.400	(385)
-reference group invalids	1.189	(48)	1.441	(48)
R^2	0.1908		0.1614	