Journal of Health Economics 29 (2010) 426–437

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

A limited-sample benchmark approach to assess and improve the performance of risk equalization models

Pieter J.A. Stam[a,b,*], René C.J.A. van Vliet[a], Wynand P.M.M. van de Ven[a]

[a] Institute of Health Policy and Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands
[b] SiRM – Strategies in Regulated Markets, P.O. Box 24355, 3007 DJ Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 18 April 2008
Received in revised form 26 January 2010
Accepted 8 February 2010
Available online 17 February 2010

JEL classification:
G22
I10
I11
I18

Keywords:
Health insurance
Managed competition
Subsidies
Risk equalization
Optimal risk adjustment

ABSTRACT

A new method is proposed to assess and improve the performance of risk equalization models in competitive markets for individual health insurance, where compensation is intended for variation in observed expenditures due to so-called S(ubsidy)-type risk factors but not for variation due to other, so-called N(on-subsidy)-type risk factors. Given the availability of a rich subsample of individuals for which normative expenditures, Y^{NORM}, can be accurately determined, we make two contributions: (a) any risk equalization scheme applied to the entire population, Y^{REF}, should be evaluated through its performance in the subsample, by comparing Y^{REF} with Y^{NORM} (not by comparing Y^{REF} with observed expenditures, Y, in the entire population, as commonly done); (b) conventional risk equalization schemes can be improved by the subsample regression of Y^{NORM}, rather than Y, on the risk adjusters that are observable in the entire population. This new method is illustrated by an application to the 2004 Dutch risk equalization model.

1. Introduction

In several countries, competition among health insurers is used to stimulate efficiency and responsiveness to consumers' preferences in the health care sector.[1] The ultimate goal is to stimulate health insurance companies to act as prudent purchasers or providers of care for their members. At the same time, financial transfers are needed in such markets for individual health insurance in order to avoid problems of access to coverage for those at high risk. The first and best solution in this case is to organize a system of risk-adjusted equalization payments (Van de Ven et al., 2000) distributed by a sponsor via a so-called Risk Equalization Fund (REF). In European countries, the role of the sponsor is played by the government. Risk adjustment is usually based on a regression model relating observed expenditures to risk factors.

Variation in observed health care expenditures will be determined by various risk factors, not all of which the sponsor may want to subsidize. In general, the total set of potential risk factors can therefore be divided into two categories: the subset of risk factors that cause variation in expenditures which the sponsor decides to subsidize, the S(ubsidy)-type risk factors, and the subset that causes variation in expenditures which the sponsor does not want to subsidize, the N(on-subsidy)-type risk factors (Van de Ven and Ellis, 2000, pp. 768–769). In most countries, up to a certain extent, gender, health status, and age will probably be considered as S-type risk factors. Examples of potential N-type risk factors are a high propensity for medical consumption, living in a region with high prices and/or overcapacity resulting in supply-induced demand, or using providers with an inefficient practice-style (Van de Ven et al., 2000). The selection of S-type risk factors plays a crucial role in the scientific and political debate.[2] Ultimately, if government is the sponsor, this

* Corresponding author at: Institute of Health Policy and Management, Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. Tel.: +31 10 4088584; fax: +31 10 4089094.
E-mail address: piet.stam@sirm.nl (P.J.A. Stam).

[1] For example, Belgium, Germany, Switzerland, The Netherlands and USA (Van de Ven and Ellis, 2000).

[2] The need for a society to make its goals explicit is in accordance with the WHO recommendations to make societal goals for countries explicit (Murray and Frenk, 2000).

0167-6296/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhealeco.2010.02.001

categorization will be determined by value judgments in society.

Given a specific categorization of S-type and N-type risk factors, adequate measures thereof should be found in order to be able to implement a system of risk-adjusted equalization payments. However, although it may be relatively easy to collect information on age and gender, it often proves difficult to find direct measures of health status that can be made available for every insured individual. As a consequence, a rather limited set of indirect health status measures is often used instead, which may not only lead to undercompensation for expenditure variation caused by S-type risk factors, but also increases the risk of undesired compensation for expenditure variation caused by N-type risk factors. For example, working status may be used as an indirect measure of health status, although expenditure differences between employees and self-employed people may be partly caused by an N-type risk factor such as time price ("no time to visit a doctor") and the resulting propensity for visiting a doctor. Although there are econometric techniques to avoid compensation for N-type variation, such an approach is seldomly applied because it often turns out to be an even bigger challenge to find adequate measures of the N-type risk factors for every insured individual than finding adequate measures of the S-type risk factors. To the extent that the sponsor might find (more) precise measures of S-type and N-type risk factors for a limited subsample of insured people, up till now a methodology was lacking to exploit this additional information to improve the equalization payments for the total population.

The conventional method to determine the performance of a risk equalization model is to compare model predicted expenditures to observed expenditures. This study introduces a new method to assess the extent to which a given set of risk adjusters generate risk-adjusted equalization payments as intended by the sponsor. The basic idea of this new method is to develop a comprehensive "risk equalization" model for a subsample of insured people for whom more (precise) measures of S-type and/or N-type risk factors can be collected (for example, from tailor-made health surveys) than for the larger population of insured people on which the risk equalization model is estimated by the sponsor. We claim that performance of any risk equalization model should be assessed in the subsample as the difference between predicted expenditures from the former model (instead of observed expenditures) and the latter model predicted expenditures.

A second innovation is that the results of this exercise can also be used to improve the performance of conventional risk equalization models, even though the extensive array of direct measures of the S-type and N-type risk factors are available for those in the limited subsample alone. To this end, predicted expenditures from the comprehensive model in the subsample (instead of observed expenditures) should be regressed on the risk adjusters that are observable in the entire population. The approach proposed in this study is relevant for all sponsors who need to assess and improve the extent to which their system of risk-adjusted equalization payments functions in accordance with their policy goals.

In Section 2 the conceptual framework of our new approach is described. Section 3 gives an illustration by applying this method to the 2004 Dutch risk equalization model. Section 4 concludes, and Section 5 discusses the results.

2. Method

2.1. The calculation of the normative expenditures

Theoretically, the calculation of the risk-adjusted equalization payments should be based on acceptable costs, i.e. the costs of services that follow from a quality, intensity and (demand and supply) price level of treatment that the sponsor considers to be acceptable to be subsidized (Van de Ven and Ellis, 2000). In practice, however, such costs are hard to determine and therefore acceptable costs are usually based on observed expenditures instead of need-based costs. We follow this convention in this study. The calculation of acceptable costs should then be based on a prediction model of observed expenditures that includes (current or prior year) measures of both S-type and N-type risk factors as explanatory variables. This prediction model is denoted as follows:

$$Y = \beta_0 + \sum_{l=1}^{L} \beta_l S_l + \sum_{m=1}^{M} \gamma_m N_m + \xi \tag{1a}$$

where Y is the health care expenditures observed during some period in time, S_l is the lth S-type adjuster, $l = 1, ..., L$, N_m is the mth N-type adjuster, $m = 1, ..., M$, and ξ_t is an independent and identically distributed error term. The S-type and N-type adjusters may be observed prior to the observation period for expenditures or during the same period, i.e. the risk equalization model may be either prospective or concurrent, respectively. The variables Y, S_l, N_m, and ξ are $N \times 1$ vectors, the elements of which contain the observations with respect to insured individuals $i = 1, ..., N$. The α_j coefficients may vary over time.

After the estimation of the coefficients β and γ by ordinary least-squares (OLS), the acceptable costs are approximated by

$$Y^{NORM} = \hat{\beta}_0 + \sum\nolimits_{l=1}^{L} \hat{\beta}_l S_l + \sum\nolimits_{m=l}^{M} \hat{\gamma}_m \bar{N}_m \tag{1b}$$

where the values of the N-type adjusters are set equal to some level desired by (or: 'acceptable' to) the sponsor according to the so-called Schokkaert approach, and is often set equal to the overall sample mean $\bar{N}_m$ (Carr-Hill et al., 1994; Schokkaert et al., 1998; Schokkaert and Van de Voorde, 2000, 2004). Acceptable costs derived this way are called normative expenditures in this study. This Y^{NORM} is the definition of normative expenditures as good as the sponsor can get at the time of the construction of model equation (1b), given the available measurement set of S-type and N-type risk factors.

It often turns out to be quite a challenge to find adequate measures of the S-type and N-type risk factors for every insured individual. The difficulty to find precise measures of S-type and N-type risk factors arises because these measures should ideally satisfy the criteria of fair payments, appropriate incentives, and feasible data (Van de Ven and Ellis, 2000) for every insured individual in the population. For example, the latter criterion requires that the measures are feasible to obtain for all individuals without undue expenditures of time or money. As a consequence, the measurement set reduces the larger the population of insured people for which the individual risk-adjusted equalization payments should be calculated. Usually, estimation is based on as large a "sample" of individuals as possible, ideally the total population of insured people. In the context of the Dutch REF this amounts to (nearly) the entire population of over 16 million people; for the Medicare system in the USA a sample of 5% of the relevant population is used, amounting to several millions. As a consequence, the calculation of the normative expenditures is often based on a limited set of potentially inaccurate measures of the S-type and N-type risk factors.

2.2. The calculation of the REF predicted expenditures

In most countries with a system of risk equalization, an approximation of normative expenditures follows from the estimation of