TI 2013-017/V
Tinbergen Institute Discussion Paper

Can Risk Adjustment prevent Risk Selection in a Competitive Long-Term Care Insurance Market?

Piet Bakx[1]
Erik Schut[1]
Eddy van Doorslaer[1,2]

[1] *iBMG, Erasmus University Rotterdam;*
[2] *Erasmus School of Economics, Tinbergen Institute.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Can risk adjustment prevent risk selection in a competitive long-term care insurance market?

Pieter Bakx[a*], Frederik Schut [a], Eddy van Doorslaer[a,b,c]

[a] Institute of Health Policy and Management, Erasmus University, P.O. Box 1738 3000DR Rotterdam, The Netherlands.

[b] Erasmus School of Economics, Department of Applied Economics, Erasmus University, P.O. Box 1738, 3000DR Rotterdam, The Netherlands

[c] Tinbergen Institute, P.O. Box 1738, 3000DR Rotterdam, The Netherlands

* Corresponding author at: Institute of Health Policy and Management, Erasmus University, P.O. Box 1738 3000DR Rotterdam, The Netherlands. tel. +31 10 4082495, fax +31 10 4089081

E-mail addresses: bakx@bmg.eur.nl (P Bakx), vandoorslaer@ese.eur.nl (E van Doorslaer), schut@bmg.eur.nl (F Schut)

Keywords: risk adjustment, long-term care, managed competition, public insurance

JEL Code: H51, I11, I13, I18, L13

Number of words in abstract: 150

Number of words in main text: 5905

Number of tables/figures: 2/1

Acknowledgements

The authors would like to thank Statistics Netherlands for (remote) access to the linked datasets, Claudine de Meijer for sharing her Stata code which has been used to prepare the data and Maria Trottmann, Richard van Kleef, Wynand van de Ven and René van Vliet and participants at the Risk Adjustment Network conference and the 2012 European Conference on Health Economics for comments. This study is part of the Netspar theme Health and Income, Work and Care across the Life-Cycle II.

Highlights

- Competing public long-term care (LTC) insurers have incentives for risk selection
- Risk adjustment in LTC insurance is necessary to prevent selection
- Risk adjustment may prevent selection based on prior use of LTC and medical care may prevent selection
- Incentives for risk selection against some easily identifiable subgroups persist

Abstract

When public long-term care (LTC) insurance is provided by insurers, they typically lack incentives for purchasing cost-effective LTC. Providing insurers with appropriate incentives for efficiency without jeopardizing access for high-risk individuals requires, among other things, an adequate system of risk adjustment. While risk adjustment is now widely adopted in health insurance, it is unclear whether adequate risk adjustment is feasible for LTC because of its specific features. We examine the feasibility of risk adjustment for LTC insurance using a rich set of linked nationwide Dutch administrative data. Prior LTC use and demographic information are found to explain much of the variation, while prior health care expenditures are important in reducing predicted losses for subgroups of health care users. Nevertheless, incentives for risk selection against some easily identifiable subgroups persist. Moreover, using prior utilization and expenditure as risk adjusters dilutes incentives for efficiency, but using multiyear data may reduce this disadvantage.

1. Introduction

All around the world, health policy makers are confronted with ageing populations and rising demands for long-term care (LTC) and are looking for ways to guarantee access to LTC services in a sustainable way. Barr (2010) argues there is a strong case for public provision of LTC insurance. And indeed, virtually all OECD countries have at least some publicly provided mandatory coverage against LTC expenditures. Several of these countries (*e.g.* Belgium, Switzerland and the US Medicare and Medicaid programs) have integrated some “medical” LTC services in their public health insurance schemes. Other countries, *e.g.* the Netherlands (since 1968), Germany (since 1995), Japan (since 2000) and South-Korea (since 2002), have a separate public LTC insurance scheme, Typically, if public LTC insurance is provided by public non-competing insurers, these agents are not at risk for the LTC expenses of their enrollees. As a consequence, these public insurers have no incentives for securing cost-effective provision of LTC services. Instead, governments have traditionally relied on demand rationing (*e.g.* copayments, coverage restrictions) and supply rationing (*e.g.* price regulation, provider budgets) to control expenditures in public LTC insurance (Costa-Font and Courbage 2012). Both types of rationing, however, have important drawbacks, which are likely to be exacerbated by the expected increase in demand for LTC. Demand-side rationing may result in access problems for low-income individuals who need LTC; supply-side rationing may result in waiting lists and substandard quality of care.

An alternative or complementary strategy to increase the future sustainability of public LTC insurance is to provide public LTC insurers with incentives to foster efficiency and to control costs. To achieve this, public LTC insurers are made financially accountable for the LTC expenses of their enrollees and consumers are empowered to choose the most efficient LTC insurer. Introducing consumer choice and financial risk for LTC insurers, however, is likely to result in premium differentiation or, if for equity reasons risk rated premiums are not allowed, risk selection. Consumer choice and financial risk may thus reduce horizontal equity in financing and access to LTC services. Furthermore, this situation may lead to welfare losses resulting from resources employed for risk selection rather than for improving care, result in inefficient health plans that nonetheless survive and bring about lower-than-desired quality of care (van de Ven and Ellis 2000).

These negative side effects may be mitigated by risk-adjusted premium subsidies that ensure that each potential enrollee is equally attractive for insurers[1]. The combination of risk-adjusted capitation payments, community-rated premiums, and open enrollment to ensure competition among risk-bearing insurers in social health insurance markets, is known as managed (or regulated) competition (Enthoven 1988, Van de Ven and Ellis 2000). Managed competition has become widespread in health insurance to balance efficiency and equity goals and knowledge on the prerequisites for its implementation in health insurance has been accumulated. .But managed competition has not been experimented with in LTC insurance yet. Only in countries where some (medical) LTC services are integrated within health insurance, *i.e.*

[1] A system of risk-adjusted capitation payments will also be necessary if LTC insurance is carried out by regional single payers that are at risk for providing LTC to a defined population, *e.g.* regions/municipalities. But in this alternative scenario, a lower level of precision in predicting individual LTC expenses is required because risk selection is less relevant. Hence, the primary goal of these subsidies would then be to foster equity across regions or municipalities.

Belgium, Switzerland and the United States (Medicare), there is some experience with managed competition in LTC.

While risk adjustment in health insurance has been studied extensively, empirical research on risk selection and risk adjustment in LTC insurance is nearly nonexistent. The aim of this paper is to examine how and to what extent a system of ex-ante risk adjustment subsidies can reduce the scope for risk selection in LTC insurance.[2] The following five questions are addressed: (1) How do LTC expenditures differ from expenditures on medical care and how do these differences affect the options to use risk adjustment to reduce risk selection? (2) What are the predicted losses and gains on LTC in case of community rating and without risk adjustment? (3) To what extent are the predicted losses and gains reduced by the most comprehensive risk-adjustment model based on the available administrative data? (4) What is the contribution of i) demographics, ii) prior LTC use and iii) prior health care expenditures and inpatient hospital diagnoses to the reduction of the predicted losses and gains in the most comprehensive risk-adjustment model? (5) How are the predicted losses and gains affected when the risk adjusters that provide substantial perverse incentives from the risk adjustment model are removed?

[2] Definitions of LTC vary internationally. In this paper we focus on elderly care, which in the Netherlands (and elsewhere) accounts for the majority of total LTC expenditure covered by LTC insurance (CVZ 2011). Elderly care is defined as home care, social assistance, assistance with activities of daily living and inpatient stays in either a residential home or a nursing home. This definition comprises both "medical" and "non-medical" LTC: unlike in some other European countries, in the Netherlands there is no sharp distinction between medical and nonmedical LTC.

2. What is already known?

Unlike in LTC insurance, in health insurance risk adjustment is widespread. But the experience from risk adjustment in health insurance cannot be readily used to develop an appropriate risk adjustment system for LTC insurance. LTC expenditures differ from health care expenditures (HCE) in at least two important aspects (Van de Ven 2005). First, LTC expenditures are concentrated among a limited group of beneficiaries, and, conditional upon use, high and stable over time. These characteristics make risk selection based on prior expenditures much easier in LTC insurance than in health insurance in the absence of risk adjustment. Second, LTC expenditures do not only depend on the patient's health and disabilities but also on the availability of informal care (see *e.g.* Bonsang 2009). The availability of informal care is difficult to quantify with administrative data and hence cannot be fully captured by the risk adjustment formula.

Little is known about how these issues can be dealt with and about how to design appropriate risk adjustment for LTC insurance.. To date, there is only one study about the feasibility of risk adjustment in Dutch LTC insurance. Using prior LTC expenditure data from one sickness fund as a risk adjuster, Van Barneveld *et al.* (1997) examine the remaining potential for risk selection in the Dutch public LTC insurance scheme. They find an R^2-statistic of 0.90, which indicates that LTC expenditures are highly predictable at the individual level when information on prior expenditures is available. However, using prior expenditures as a risk adjuster means that the insurer will be compensated (partly) for high(er) expenditures through higher future risk-adjusted capitation payments. This compensation is a major drawback because it lowers incentives for efficiency compared to the situation of capitation payments that are not based on prior expenditures. Still, risk adjustment based on prior expenditures (or use) improves incentives for efficiency compared to the situation in which insurers are fully reimbursed for all costs if insurers compete on quality (Marchand *et al.* 2003).

Several studies on risk adjustment in US Medicare have tackled similar issues. Medicare covers health care, rehabilitation care in a nursing home, and some home care services for the elderly and the disabled as part of a broader package of health insurance (CMS 2011). While the Medicare benefit package encompasses far more than just LTC insurance, the Medicare target group is similar and studies on risk adjustment in Medicare still provide a number of relevant insights. First, risk adjustment for private insurers that offer Medicare Advantage plans and receive a capitation payment for each enrollee is extensive and takes into account frailty as measured by the number Activities of Daily Living (ADL) problems for the Program of All-Inclusive Care for the Elderly (PACE). The risk adjustment model without frailty is found to systematically underestimate expenditures for the frail elderly and might induce risk selection against this group (Kautter *et al.* 2009).

Second, the relationship between health care use in the past, demographic characteristics and future health expenditures changes upon institutionalization and is different for those who were eligible for Medicare because of reaching the age of 65 and those who were eligible because they were disabled (Pope *et al.* 2004). This finding implies that risk adjusters should be interacted with institutionalization and age to reflect nonlinearity.

Third, despite extensive risk adjustment, incentives for risk selection persist: while risk selection on expected costs decreased after expanding the risk adjustment formula beyond age and gender, insurers now select profitable enrollees by focusing on characteristics not included in the model. Consequently, the Medicare program has become more expensive and expenditures are transferred from those in bad health to those in good health (Brown *et al.* 2011).

Beyond the US, experience with risk adjustment in LTC insurance is mostly limited to Switzerland and Belgium. In these countries, medical LTC is included in social health insurance. In Switzerland, the risk adjustment formula comprises age, gender and a dummy variable accounting for a recent stay of at least three days in a hospital or a LTC-facility (Von Wyl and Beck 2012). This dummy variable is likely to pick up some of the variation in expected LTC expenditures. The Belgian risk adjustment formula includes more specific risk adjusters for LTC. The capitation payment is adjusted for receipt of certain allowances (*e.g.* for handicapped or because of a need for assistance) or nursing care at home during 3 months (category B or C on the Katz-scale (Katz and Akpom 1976)). The formula also includes a number of indicators related to LTC use, *e.g.* living alone, being widow/widower, physiotherapy for a severe illness, and Parkinson's disease (Schokkaert and van de Voorde 2011). While the Belgian risk adjustment formula is more sophisticated than the one used in Switzerland, the financial risk for Belgian health insurers is much more limited (Schokkaert and van de Voorde 2011; Paolucci *et al.* 2007). So, incentives for risk selection against LTC patients appear to be large in Switzerland and very limited in Belgium. It is, however, unclear whether the more sophisticated Belgian model would suffice to prevent risk selection if financial risk for insurers were expanded.

3. Data and methods

3.1 Data

We use information from four nationwide administrative datasets and one survey which are all linked by Statistics Netherlands at the individual level (see de Meijer *et al.* (2011) for a more detailed description of the data). The administrative data include (1) health care expenditures in 2000-2004 from the health insurance data collected by Vektis; (2) LTC use in 2004 and 2005 from the Central Administration Office of the LTC insurance scheme (CAK); (3) hospital admissions in 2002, 2003 and 2004 from the hospital discharge register (LMR); (4) demographic information for 2004 from the municipal register (GBA) and (5) mortality from the cause-of-death registry (CBS). In addition, the General Survey of Living Conditions (POLS) held in 2004 provides details on health, disability, and other individual characteristics for a randomly drawn, representative sample of the non-institutionalized population. The administrative data sets comprise the entire Dutch population, except for prior health care expenditures which are registered for sickness fund enrollees only (two-thirds of the population)[3] and except for LTC use, which is registered for adults only (≥ 18 years of age). Furthermore, the records for one third of those eligible for sickness fund coverage cannot not be linked; item non-response on other variables was small (1.7% of the sample) and in all cases the result of missing co-residence status. As a result, the final sample consists of individuals who were insured through a sickness fund, did not die in 2004 and whose records could be linked to the municipality register. The total study population was 5,719,934, which is 45% of the Dutch adult population in 2004.

---

[3] Until 2006 enrollment was mandatory for two thirds of the population with an income below a threshold; the remainder of the population was not eligible for social health insurance and could buy private insurance. By contrast, public LTC insurance was (and is) mandatory for the entire population.

From this subset of the population, 7790 individuals were included in the 2004 POLS survey; 3619 of these respondents also completed the more specific health module.

3.2 Methods

We identify the extent to which a risk adjustment model can reduce incentives for risk selection in three steps. First, we identify the insurers' incentives to select against subgroups based on individual characteristics in the presence of community rating but in the absence of risk adjustment. In order to quantify the insurers' incentives for risk selection, we calculate the difference between the average actual costs by subgroup and the average costs of the entire population in 2005. We consider the incentives for risk selection to be strong when the number of users in the subgroup is substantial (> 300), the predicted loss for this group – the difference between observed and predicted expenditures – is large (> 1000 euro) and the predicted loss was significantly ($p < 0.05$) different from zero. When these criteria are met, the subgroup is included in the risk-adjustment model.

Second, we build the full risk adjustment model in a stepwise manner to examine to what extent each set of individual characteristics contributes to explaining individual variation in LTC use. We first test the impact of a basic model based on demographic characteristics on the predicted loss for all subgroups. Next, we add subgroups based on (i) prior LTC use, and (ii) prior health care expenditures and hospital admissions to this basic model (section 4.3 elaborates on these variables). The full model includes all subgroups that were identified in the first model.

Third, for each subgroup that is included in the full model, we assess the impact of including this subgroup in the risk adjustment formula on the insurers' incentives for efficiency – an additional commonly used selection criterion (see *e.g.* van Kleef and van Vliet (2010), Van de Ven and Ellis (2000) and Pope *et al.* (2000)). Subgroups that have a negative impact on the insurers' incentives for efficiency are removed in the incentive compatible risk adjustment model. This third step thus sheds light on the tradeoff between creating incentives for efficiency and incentives for risk selection. Whether the predicted loss is stable over time – another common criterion for inclusion – cannot be assessed with the data set at hand.

All models described above are estimated by ordinary least squares regression (OLS) in order to facilitate interpretation of the results (Van de Ven and Ellis 2000)[4]. Moreover, all current Dutch risk adjustment models use OLS, so using OLS increases the comparability and compatibility of these models.

The subgroups based on detailed information on health status, disability and socio-economic status from the POLS survey are too small to be included in the risk adjustment model. Instead, these subgroups are used as a benchmark to evaluate the impact of the risk adjustment model on incentives for risk selection.

3.3 Variables

In each model the dependent variable measures public LTC expenditures in 2005. In case the individual dies in 2005, expenditures are annualized by dividing expenditures by the share of the year the individual was alive. The data set provides information on

[4] Other commonly used specifications did not provide a strictly better fit than OLS. Results for the other specifications are available from the corresponding author.

the quantity of LTC that was provided in kind, which is 95% (2006) of the publicly financed LTC in the Netherlands[5] (CVZ 2011). The quantities provided, i.e. days institutionalized or hours of home care, are multiplied by the maximum prices as set by the government in order to calculate expenditures; co-payments are not taken into account. The data contains information on institutional care for 2004 and 2005 and for all six types of home care for 2004. For 2005, only four out of six types of home care were in the data[6]. Overlapping stays in LTC institutions are only adjusted when both stays were in the same institution.

The set of subgroups that make up the basic model are based on three demographic characteristics: age, gender and co-residence, i.e. whether someone lived in a single-person household. Age and gender are the backbone of any risk-adjustment model, while co-residence proxies informal care availability. Informal care availability is an element of the eligibility assessment procedure for homecare (CIZ 2005) and formal LTC use is known to be correlated with informal LTC use (Bonsang 2009; Van Houtven and Norton 2004).

The subgroups of LTC users are based on prior LTC *use* rather than *expenditures* because using prior LTC use as a risk adjuster still rewards insurers for negotiating lower prices with providers. Subgroups are created for each type of home care and each type of institutional care separately. Each of the subgroups of home care users consists of individuals who used this specific type of home care at least one hour per week on average. In selecting subgroups of institutional care users, we aim at

[5] The remaining 5% consisted of cash transfers, which are not in the dataset.
[6] Information on assistance (*activerende begeleiding*) and support (*ondersteunende begeleiding*) is not available.

balancing responsiveness to changes in LTC use against incentives for overreporting and oversupply resulting from the (partial) reimbursement in the future of additional expenditures. Therefore, for each of the four types of institutional care, four subgroups are generated consisting of individuals who stayed in an LTC institution for ≥ 1 day, 91-180 days, 181-365 days, and the entire year (366 days), respectively. These subgroups reflect differences in expected future costs between long-term and short-term residents: future costs are increasing in the number of days that the individual is institutionalized. Furthermore, following Van Barneveld *et al.* (1997), two subgroups are created consisting of patients who received home care and institutional care, respectively, on the last day of 2004.

Subgroups are included for five specific HCE categories that may be related to LTC use: hospital and outpatient expenditures, prescription drugs, paramedical care, transportation, and durable medical equipment. For each of these categories, three subgroups are constructed: high expenditures (being in top 15%) during the last year (omitted for hospital and outpatient expenditures), during each of the last three years, and during each of the last five years. Because the data only includes HCE covered by sickness funds, we also include a variable indicating which persons were not insured through a sickness fund in one of the four years preceding 2004. If someone was no longer registered with a sickness fund during a year (*e.g.* because of losing his/her eligibility status due to exceeding the income threshold) and hence is not in the data set for the entire year, expenditures are annualized.

In addition to the subgroups based on prior HCE, we also create subgroups based on hospital admissions because information on hospitalization and diagnosis information may help to predict LTC use (Wong *et al.* 2010). Subgroups are based on 94 diagnoses (based on a grouping algorithm of ICD-9 codes, see Polder *et al.* 2002) and on 48 types of treatments (based on ICD-9-CM volume 3 codes) using hospital admission data from 2002-2004. These data are also used to create 12 Diagnostic Cost Groups (DCGs). DCGs are used for risk adjustment in the Dutch health insurance scheme and consist of clinically homogenous inpatient diagnoses for chronic health problems (*e.g.* cancer and COPD) that have similar future HCE (Van de Ven and Ellis 2000). Each individual is assigned to either the reference group (DCG 0) - people with no hospital admission or an incidental admission (*e.g.* fractures) - or the highest DCG (corresponding to the highest expected costs) that they are eligible for. Prinsze and Van Vliet (2007) provide an extensive description of DCGs in the Netherlands and the Health Insurance Regulation (*Regulering van de zorgverzekering*) provides the classification of diagnoses into DCGs (Rijksoverheid 2005)[7] using the ICD-code of the main diagnosis and the medical specialty that set this diagnosis. This classification is then used to assign individuals to DCGs. Because the subgroups based on diagnosis and treatments and the DCGs overlap, we include the DCGs but not the separate subgroups based on diagnoses and treatments in the risk adjustment model. Furthermore, the impact of the DCGs on the incentives for efficiency is known to be limited while including all subgroups separately will provide incentives for oversupply and overreporting.

---

[7] The assigned DCG does not match with the actual DCG for some individuals because of two limitations of the data set: 1. not all hospitals reported information on patients to the national medical registry; 2.information on two relevant ‘side treatments’, dialysis and artificial respiration at home was not available at all. As a consequence, DCG 13 (dialysis) is empty and the reference category consists of DCG0, DCG13 and patients who needed artificial respiration and should therefore be in DCG12. Furthermore, information on radiotherapy and chemotherapy was not specific enough to ensure that no patients who do not belong in the related DCG are excluded.

As the administrative data do not provide detailed information on personal characteristics, subgroups based on health, disability and socio-economic characteristics could only be retrieved from the smaller set of respondents that completed the POLS survey. Although it is much smaller and persons in nursing homes are not sampled, this survey information is important because it allows investigating incentives for insurers to use such questionnaires for risk selection purposes. The same subgroups are used as in De Meijer *et al.* (2011), who study determinants of LTC expenditures among the elderly, and in Stam and Van de Ven (2008), who identify subgroups that generate losses for health insurers. Of these subgroups, only those are selected for which the predicted loss deviates significantly from zero in the absence of risk adjustment. Because the average predicted profit without risk adjustment for the POLS sample and the subsample answering the health module are positive, the predictions for these samples are adjusted by subtracting the mean deviation from zero for the relevant sample multiplied by the ratio of the individual's observed expenditures and the sample mean observed expenditures to ensure that the average predicted profit was zero.

4. Results

4.1 Descriptive statistics

Figure 1 and table 1 show that the distribution of LTC costs within the population is highly skewed. The median is at 4,598 euro; 2 out of 3 LTC users spent less than 10,000 euro. Furthermore, there are two spikes, one at 32,000 euro (a full year of care in a residential home) and one at approximately 91,000 euro (a full year of care in a nursing home). The average cost per LTC user (15,677 euro) is much higher than the average cost per user of medical care (about 2000 euro in 2004). Furthermore, LTC expenditures are strongly correlated with prior use of LTC: average LTC expenditures in 2005 are higher for home care users in 2004 than for non-users and highest for nursing home residents in 2004 (table 1). This finding confirms that LTC expenditures are quite stable over time. Table 2 shows the subgroups derived from administrative data that are considered for the risk adjustment formula as well as the number of individuals in each subgroup.

[Figure 1]

[Table 1]

4.2 Analysis

The regression analysis reveals that the included covariates explain a large share of the variation in aggregate costs of LTC use in 2005: the R^2-statistics are generally higher than those obtained in similar studies on medical care and mental health care (see *e.g.* van de Ven and Ellis 2000)[8]. Most of the explanatory power derives from the demographic variables and prior LTC use. Demographics alone contribute to an R^2 of 0.23. Including prior LTC use raises the R^2-statistic to 0.73, while variables related to prior HCE contribute only marginally to the overall goodness of fit, regardless of whether prior LTC use is included. A Copas test (Copas 1983) did not detect overfitting. Nearly all coefficients are significant in each of the models and show the expected sign. The DCGs sometimes violate the monotonicity requirement: being assigned to a higher DCG with a more severe diagnosis does not in all cases lead to a

[8] Appendix 1 contains descriptive statistics and all regression results.

higher capitation payment. This is undesirable as it generates disincentives for providing more care. The most prominent example is DCG 4, which includes diagnoses related to a cardiovascular accident, myocardial infarct, and angina pectoris among other things, and which has the third largest coefficient. These results highlight that the relationship between prior hospital stays and LTC expenditures is different from the relationship between prior hospital stays and HCE, on the basis of which these DCGs were constructed.

No risk adjustment model

In case of community rated premiums but no risk adjustment, the predicted losses would be very large for subgroups based on prior LTC use or based on prior health care expenditures (table 2)[9]. These predicted losses, together with the large size of most of these subgroups (last column), signal that incentives for risk selection against these subgroups would be huge. Other results (available upon request) show that some diagnoses are indicators of a persistent loss: for four diagnoses that yield a large predicted loss in the next year, the predicted loss is still larger than 1000 euro two years later and three years later.

[Table 2]

[9] The appendix contains the predicted losses for the all subgroups that were included in the final model, 20 subgroups based on diagnosis from information on hospital admissions in 2004 and subgroups that were based on the POLS survey data; results for other subgroups are available from the corresponding author.

Demographic model

The results for the Demographic Model, which would adjust subsidies by age, gender and co-residence status of the enrollee, show that including demographic characteristics in the risk adjustment model does not sufficiently reduce the predicted losses for subgroups based on prior LTC use and prior HCE (table 2). Therefore, it seems imperative to include the latter subgroups in the risk adjustment model to reduce incentives for risk selection.

Prior LTC model

Including variables on prior LTC use as risk adjusters by definition reduces the predicted losses on these subgroups to zero. But risk adjustment based on prior LTC use not only reduces predicted losses for prior LTC users but also for many subgroups based on prior HCE. In addition, this risk-adjustment model reduces the predicted losses for several subgroups of individuals who were hospitalized for diagnoses that were associated with the highest predicted loss without risk-adjustment (table 2). However, it does not substantially reduce the predicted losses for subgroups of LTC users that are not included, *e.g.* individuals who used LTC on the last day of the year.

Prior HCE and DCG model

Subsequently, we examined the effect of adding information on prior health care use and HCE patterns in the risk adjustment formula on the predicted losses. The predicted losses for the subgroups of insured that used LTC in 2004 all remain above the threshold of 1000 euro when DCGs are added to the model, along with variables indicating high expenditures (top 15%) on hospital and outpatient care for the last three and the last five years, and high expenditures on prescription drugs, transport, and durable medical equipment for the last year, the last three and the last five years (table 2). But while these variables only have a small impact on the predicted loss for LTC users, including HCE is important for reducing the predicted loss for subgroups based on prior hospital admissions for several diagnoses, *e.g.* heart failure, and asthma and COPD. So whereas for some diagnoses, prior LTC use is more important in reducing the predicted loss, for other diagnoses prior HCE and DCGs causes the largest drop in the predicted losses.

Full model

When all information is combined in the full risk adjustment model, the predicted losses are reduced substantially for many of the subgroups we distinguished. For example, this full model reduces predicted losses sufficiently for all but seventeen diagnoses and for all but one type of treatment. Yet, including information on prior HCE and the variables on LTC use also leads to predicted profits larger than 1000 euro for two diagnoses: prostate cancer and chronic ulcers of skin including decubitus.

The initial predicted losses also vanish for the subgroups based on self-reported disability, health and socio-economic status all characteristics when prior LTC use and prior HCE are included in the risk adjustment formula. Although the loss is still larger than 1000 euro for subgroups unable able to perform at least one ADL, it is no longer significantly different from zero (table 2).

Incentive compatible model

All subgroups based on prior LTC use and listed in table 3 are large and generate a large predicted loss in the absence of risk adjustment. Some of these subgroups are nonetheless excluded from the full model because their inclusion is expected to give insurers too strong incentives to overreport and oversupply. For example, the required additional spending for admitting a person for a single day in a nursing home (about 190 euro – see appendix) is much lower than the subsequent increase in the risk-adjusted capitation payment of 11299 euro for the subgroup of people that are admitted to a nursing home for 1-90 days. Table 2 shows that when an individual uses LTC during a given year, in the next year the insurer would be compensated for most of the loss if risk adjustment were based on prior LTC utilization. In addition to affecting the insurer's incentives for efficiency, including prior expenditures also makes insurers insensitive to prices.

The trade-off between incentives for efficiency and incentives for risk selection is also relevant for some subgroups based on prior HCE and health care use. For DCGs and subgroups with high HCE in successive years the incentive problem is expected to be limited (Van de Ven and Ellis 2000). But for some subgroups, most notably individuals with high expenditures on transportation or medical equipment, the inclusion criteria are set at low levels because very few individuals use these services. As a result, for these groups the minimum amount of expenditures is lower than the increase in the risk adjustment payment. Therefore, the subgroups based on only high expenditures in the previous year are omitted in the incentive compatible model.

Leaving subgroups that were expected to compromise insurers' incentives for efficiency out of the incentive compatible model had a small effect on the overall predictive power of the model: the incentive compatible model has an R^2-statistic of 0.70, compared to 0.73 for the full model. A comparison of the results of the full model and the incentive compatible model at the subgroups level reveals that removing these risk adjusters did not only affect the predicted losses for the subgroups that were no longer included; but it also affected the predicted losses for subgroups based on hospital diagnoses and treatments and for the subgroups based on detailed survey information on health and disability. Yet, the impact on the predicted losses for these other subgroups was often fairly limited. Therefore, further reduction of the number of subgroups in the risk adjustment model may be considered.

5. Conclusion and discussion

In several countries public LTC insurance is offered by non-competing agents that are not at risk for providing coverage. This situation is suboptimal because it provides these agents with little or no incentive for efficiency and cost containment. Transforming or replacing these agents by insurers that bear the financial risk is easy but might lead to socially undesired outcomes in terms of equity and efficiency in absence of adequate risk adjustment. We have investigated the scope for risk selection and the feasibility of a LTC risk adjustment formula that sufficiently reduces incentives for risk selection. Our findings demonstrate that a model that is only based on demographic characteristics performs poorly: subgroups that may be identified based on their prior LTC use, prior HCE or other individual characteristics yield a significant predicted loss to the insurer in case of community rated premiums. Subsequently, we investigated the impact of 1) including individual-level information on prior health care and LTC use and 2) excluding risk adjusters that compromise insurers´ incentives for efficiency.

Not surprisingly, prior use of LTC services shows up as the best available predictor of future LTC use and hence its inclusion reduces incentives for risk selection substantially. Its main drawback is that it simultaneously reduces incentives for efficiency because insurers will be compensated for expenditures made in the preceding year. When more data on prior LTC use become available in the near future, the problem of reduced incentives for efficiency may at least partially be overcome by including indicators for having used (any) LTC for multiple years.

Next to prior LTC use, prior HCE and inpatient diagnosis and treatment information also prove to be vital: predicted losses persist for certain categories of HCE and for some inpatient diagnoses that occur mostly among the frail elderly even when prior LTC use is taken into account. These diagnoses probably indicate a negative health shock that causes the onset or intensification of formal LTC use.

Extending the model based on demographic characteristics with prior LTC use, prior HCE and inpatient diagnosis and treatment information does not fully eliminate the potential for risk selection. While the predicted losses disappear for health, disability and socio-economic characteristics that can be obtained from a survey, risk selection on the basis of some inpatient diagnoses and treatments and prior LTC use in the risk adjustment model remains feasible. Most of these subgroups are easy to identify for an insurer, *e.g.* patients who received short-term institutional LTC, were admitted to a hospital for a hip fracture, dementia-related problems or asthma or COPD, or who had high HCE in 2004 but not in 2003 or 2002. Yet, including these variables in the risk adjustment formula is not an option, as it would give insurers an incentive for overprovision. This problem may partly be overcome by recalibrating the DCG for predicting LTC expenditures.

Ideally, risk adjustment is based on data on individuals' underlying needs for care. But this information is usually not included in administrative data and insurers' LTC claims data. As a consequence, in the Netherlands as elsewhere, risk adjustment in LTC will have to rely on prior utilization and expenditure data, which is likely to not only reduce incentives for risk selection but also for efficiency. The risk adjustment model derived in this paper might nonetheless still be improved further, even with existing data. Incentives for risk selection may be reduced further by recalibrating the DCGs and by including more subgroups, *e.g.* based on socio-economic status and more specific information on prior use of durable medical equipment that indicates disability (see *e.g.* Van Kleef and Van Vliet 2011) as risk adjusters. In addition, using data for multiple prior years will also substantially reduce the problem of creating incentives for overreporting and oversupply. Most of all, our findings highlight the interrelatedness of elderly care, medical care and social care. Our findings show that to prevent risk selection, any risk adjustment formula needs to take into account the potential simultaneous or subsequent use of these related types of care. Consequently, these findings also have implications for the reverse relationship: taking into account prior LTC use should also be considered and studied for optimizing risk adjustment in health insurance.

References

Barr N. Long-term care: a suitable case for social insurance. Social Policy & Administration 2010; 44(4); 359-374.

Bonsang E. Does informal care from children to their elderly parents substitute for formal care in Europe? Journal of Health Economics 2009; 28 (1); 143-54.

Brown J, Duggan M, Kuziemko I, Woolston W. How does risk selection respond to risk adjustment? Evidence from the Medicare Advantage program. NBER 2011; working paper 16977

CIZ. Protocol gebruikelijke zorg. CIZ: Driebergen; 2005

CMS. Medicare Coverage of Skilled Nursing Facility Care. http://medicare.gov/Publications/Pubs/pdf/10153.pdf. 2007. Accessed on 9 April 2011

Copas J. Regression, prediction and shrinkage. Journal of the Royal Statistical Society. Series B (Methodological) 1983; 45 (3); 311-354

Costa-Font J, Courbage C (eds.). Financing long-term care in Europe. Institutions, markets and models. Palgrave Macmillan: Houndmills, Basingstoke; 2012

CVZ. Zorgcijfers. http://www.cvz.nl/zorgcijfers. 2011. Accessed on: 30 September 2011

De Meijer C, Koopmanschap M, Bago d’Uva T, van Doorslaer E. Determinants of long-term care spending: Age, time to death or disability? Journal of Health Economics 2011; 30 (2); 425-38.

Enthoven A. Theory and practice of managed competition in health care finance. Amsterdam: North Holland; 1988

Katz S, Akpom C. A measure of primary sociobiological functions. International Journal of Health Services 1976; 6; 493-508.

Kautter J, Ingber M, Pope G. Medicare risk adjustment for the frail elderly. Health Care Financing Review 2009; 30 (2); 83-93.

Marchand M, Sato M, Schokkaert E. Prior health expenditures and risk sharing with insurers competing on quality. Rand Journal of Economics 2003; 34 (4); 647-669.

Paolucci F, Schut F, Beck K, Van de Voorde C, Gress S, Zmora I. Supplementary health insurance as a tool for risk selection in mandatory basic health insurance markets. Health Economics, Policy and Law 2007; 2(2); 173-192

Pope G, Ellis R, Ash A, Liu C, Ayanian J, Bates D, Burstin H, Iezzoni L, Ingber M. Principal inpatient diagnostic cost group model for medicare risk adjustment. Health Care Financing Review 2000; 21 (3); 93-118.

Polder J, Takkern J, Meerding W, Kommer G, Stokx L. *Kosten van Ziekten in Nederland. De zorgeuro ontrafeld*. RIVM: Bilthoven; 2002

Prinsze F, van Vliet R. Health-based risk adjustment: improving the pharmacy-based cost group model by adding diagnostic cost groups. Inquiry 2007; 44(4); 469-480

Rijksoverheid. *Bijlage 10 van de Regeling van de zorgverzekering*. http://wetten.overheid.nl/BWBR0018715/Bijlage10. 2005; Accessed on 15 April 2011

Schokkaert E, Van de Voorde C. Belgium's health care system: should the communities/regions take it over? Or the sickness funds? Re-Bel e-book 10, http://www.rethinkingbelgium.eu/; 2011

Stam P, van de Ven W. De harde kern van de risicoverevening. Economisch Statistische Berichten 2008; 93 (4529); 104-107.

Van Barneveld E, van Vliet R, van de Ven W. Risk-adjusted capitation payments for catastrophic risks based on multi-year prior costs. Health Policy 1997; 39; 123-135

Van de Ven W. *Second opinion "risicoverevening AWBZ"*. Erasmus University: Rotterdam; 2005

Van de Ven W, Ellis R 2000. Risk adjustment in competitive health plan markets. In: Culyer A, Newhouse J (Eds). Handbook of health economics, volume 1a. North Holland: Amsterdam; 2000 p. 755-845..

Van Houtven C, Norton E. Informal care and health care use of older adults. Journal of Health Economics 2004; 23; 1159-1180.

Van Kleef R, van Vliet R. Prior use of durable medical equipment as a risk adjuster for health-based capitation. Inquiry 2010; 47 (4); 343-358.

Van Oyen H, Van der Heyden J, Perenboom R, Jagger C. Monitoring population disability: evaluation of a new Global Activity Limitation Indicator (GALI). International Journal of Public Health 2006; 51; 153-161

Von Wyl V, Beck K. 2012. How to measure risk selection in a competitive health insurance market? *Mimeo*

Ware J, Kosinski M, Keller S. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Medical Care 1996; 34 (3); 220-233

Wong A, Elderkamp-de Groot R, Polder J, van Exel J. Predictors of long-term care utilization by Dutch hospital patients aged 65+. BMC Health Services Research 2010; 10: 110.

Figure 1: Distribution of LTC expenditures in 2005 of LTC users in 2005

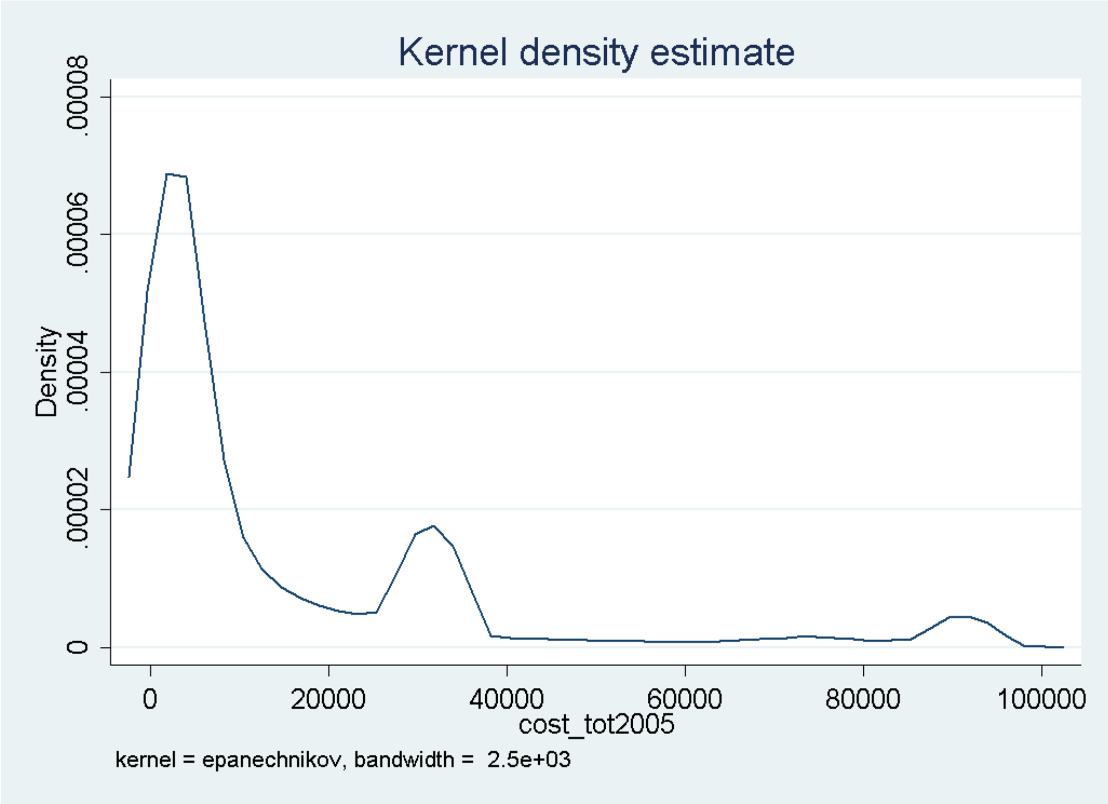

Table 1: Descriptive statistics population

	Mean	*St.dev.*
LTC expenditures in 2005	1159.18	7564.05
LTC expenditures in 2005 conditional on any use	15677.04	23370.78
LTC expenditures in 2005 if no LTC was received in 2004	93.62	1657.39
LTC expenditures in 2005 if received home care in 2004	9673.22	18333.23
LTC expenditures in 2005 if stayed in a residential care facility in 2004	31767.09	15902.13
LTC expenditures in 2005 if stayed in a nursing home in 2004	61451.47	35674.80

Table 2: Predicted losses for selected subgroups

	No risk adjustment	Demographic model	Prior LTC model#	Prior HCE and DCG model#	Full model #	Incentive compatible model#	Subgroup size
Demographic information	No	Yes	Yes	Yes	Yes	Yes	
Prior LTC use	No	No	Yes	No	Yes	Yes	
Prior HCE and diagnosis information	No	No	No	Yes	Yes	Yes	
Subgroups of LTC users in prior year							
Personal care	16992*	8293*		7269*			82116
Nursing	25799*	17183*		15303*			24727
Nursing home: combined care 1-90 days	31009*	23310*		20888*		21005*	10546
Nursing home: combined care 91-180 days	51350*	42482*		39803*			3853
Nursing home: combined care 180-365 days	71408*	61349*		59290*			4359
Nursing home: combined care 366 days	85868*	74655*		74693*			12439
Receiving home care on last day of 2004	10078*	2459*		1669*		876*	172291
Subgroups based on HCE in 2000-2004							
In top 15% in prior 3 years: hospital + outpatient care	3410*	1804*	278*				137618
In top 15% in prior 5 years: hospital + outpatient care	4031*	2321*	393*				57597
Expenditures on transportation in prior year	7350*	4271*	582*			916*	201609
Expenditures on transportation in prior 3 years	9040*	5753*	353*				43615
Expenditures on transportation in prior 5 years	8565*	5347*	374*				22462
DCG 4, *e.g.* Cardiovascular accident, stroke, *angina pectoris*	10172*	6865*	1145*				11358
Subgroups based on diagnosis information from 2004 hospital admission data							
Dementia	30423*	20777*	9255*	20114*	9274*	11692*	821
Hip fracture	21225*	12029*	-950*	9167*	-1229*	3205*	6433
Chronic ulcers of skin including decubitus	13421*	8676*	-224	4514*	-1897*	-1528*	873
Stroke	10840*	7288*	1109*	2219*	327*	1271*	11998
Heart failure	10054*	3806*	858*	745*	182	637*	8147
Diabetes mellitus including diabetic complications	6478*	4633*	679*	2770*	133	231	4746
Asthma and COPD	6128*	3721*	876*	122	-78	26	8035
Subgroups from POLS health survey (n = 4619)							

Has difficulty to/cannot perform ≥ 1 ADL[d]	16221*	8918*	357	7887*	289	532	90
Cannot perform ≥ 1 ADL	23365*	13476*	1479	11906*	1409	1365	32

*Significant at the $p < 0.05$ level

#Cells are empty if variable is included in this extension of the risk adjustment model

Appendix

Table A1: Coefficients of risk classes included in the risk-adjustment model

	Prevalence	Demographic model	Prior LTC model#	Prior HCE and DCG model#	Full model	Incentive compatible model	Min. cost of prior use to qualify for subgroup
Demographic information		Yes	Yes	Yes	Yes	Yes	
Prior LTC use		No	Yes	No	Yes	Yes	
Prior HCE and diagnosis information		No	No	Yes	Yes	Yes	
Female age < 50	26.43	-346*	-49*	-307*	-46*	-73*	
Female age: 50-64	9.54	327*	30*	147*	-7	22*	
Female age: 65-69	2.70	929*	169*	576*	101*	154*	
Female age: 70-74	2.47	2125*	404*	1569*	317*	460*	
Female age: 75-79	2.26	4733*	928*	3937*	832*	1135*	
Female age: 80-84	1.89	9469*	1863*	8444*	1770*	2286*	
Female age: 85-89	1.13	17034*	3310*	15770*	3220*	3946*	
Female age: 90+	0.68	27154*	4556*	25732*	4481*	5403*	
Male age < 50	31.01	0	0	0	0	0	
Male age: 50-64	12.44	188*	47*	55*	17*	22*	
Male age: 65-69	2.96	592*	210*	313*	154*	148*	
Male age: 70-74	2.51	1256*	401*	853*	329*	370*	
Male age: 75-79	1.90	2698*	793*	2106*	700*	847*	
Male age: 80-84	1.25	5473*	1460*	4671*	1350*	1678*	
Male age: 85-89	0.58	9961*	2333*	8999*	2229*	2784*	
Male age: 90+	0.25	16478*	3005*	15384*	2924*	3631*	
Male living alone	10.63	239*	-361*	213*	-358*	-158*	
Female living alone	11.3	2096*	72*	1917*	68*	339*	
≥ 1 hour per week in 2004: domestic care I	0.81		336*		318*	1534*	2360.80
≥ 1 hour per week in 2004: domestic care II	1.87		2594*		2548*	3865*	733.20
≥ 1 hour per week in 2004: activating support	0.06		885*		857*	1082*	1367.60
≥ 1 hour per week in 2004: guidance	0.16		4760*		4783*	5369*	2298.40
≥ 1 hour per week in 2004: personal care	1.44		7346*		7275*	9101*	2033.20
≥ 1 hour per week in 2004: nursing	0.43		12079*		11840*	13502*	3088.80

Residential home 1-90 days	0.18	6192*		6112*		86.02
Residential home 91-180 days	0.07	11869*		11715*	15147*	7827.82
Residential home 181-365 days	0.13	20595**		20458*	27012*	15569.62
Residential home 366 days	0.78	22901*		22728*	28010*	31483.32
Nursing home: somatic care 1-90 days	0.02	11578*		11299*		188.65
Nursing home: somatic care 91-180 days	0.01	24051*		23572*	26345*	17166.87
Nursing home: somatic care 180-365 days	0.01	41468*		41025*	47521*	34145.09
Nursing home: somatic care 366 days	0.02	58101*		58089*	63727*	69044.77
Nursing home: psychogeriatric care 1-90 days	0.01	29897*		29773*		203.70
Nursing home: psychogeriatric care 91-180 days	0.01	37814*		37798*	41849*	18535.95
Nursing home: psychogeriatric care 180-365 days	0.01	54145*		54055*	59739*	36868.21
Nursing home: psychogeriatric care 366 days	0.04	63126*		63206*	68788*	74551.19
Nursing home: combined care 1-90 days	0.18	20167*		19956*		247.01
Nursing home: combined care 91-180 days	0.07	37289*		36939*	38841*	22477.55
Nursing home: combined care 180-365 days	0.08	59337*		59072*	63692*	44708.10
Nursing home: combined care 366 days	0.22	78034*		78092*	83659*	90404.23
Used home care on the last day of the year	3.01	2750*		2609*		14.10
Used institutional care on the last day of the year	2.09	6366*		6279*		86.02
In top 15% in prior year: durable medical equipment	12.12		156*	79*		1
In top 15% in prior 3 years: durable medical equipment	5.78		1471*	353*		3
In top 15% in prior 5 years: durable medical equipment	3.69		1980*	370*	469*	5
In top 15% in prior year: transportation	3.52		3780*	579*		1
In top 15% in prior 3 years: transportation	0.76		5802*	209*		3
In top 15% in prior 5 years: transportation	0.39		4649*	190*	670*	5
In top 15% in prior year: paramedical care	4.30		532*	-51*		1
In top 15% in prior 3 years: paramedical care	1.60		1064*	172*		212.77
In top 15% in prior 5 years: paramedical care	0.90		1425*	371*	471*	382.81
In top 15% in prior year: drugs	12.26		-203*	21*		588.88
In top 15% in prior 3 years: drugs	8.43		-84*	55*		1761.78
In top 15% in prior 5 years: drugs	6.66		-162*	33*	64*	2743.73
In top 15% in prior 3 years: hospital + outpatient care	2.41		69*	-4	96*	2096.22
In top 15% in prior 5 years: hospital + outpatient	1.01		109*	88*	136*	3145.02

care							
No records available for all of the last 5 years	25.37			323*	76*	62*	0
No DCG	97.60			0	0	0	
DCG 1	0.39			-883*	-1518*	-116*	
DCG 2	0.49			-861*	-81*	243*	
DCG 3	0.25			-1070*	-83*	287*	
DCG 4	0.20			5323*	986*	3373*	
DCG 5	0.31			882*	312*	953*	
DCG 6	0.19			331*	148*	689*	
DCG 7	0.13			2295*	1068*	1730*	
DCG 8	0.12			1972*	809*	1445*	
DCG 9	0.13			1772*	805*	1397*	
DCG 10	0.09			4313*	3093*	3995*	
DCG 11	0.05			4770*	3647*	4540*	
DCG 12	0.05			2154*	1635*	2327*	
Intercept		-3	6*	-207*	-36*	-18*	
R^2		0.23	0.73	0.24	0.73	0.70	
Number of observations		5,719,934	5,719,934	5,719,934	5,719,934	5,719,934	

*significantly different from zero at the $p < 0.05$ level

Table A2: predicted losses for subgroups of LTC users in prior year[*]

	No risk adjustment	Demographic model	Prior LTC model[#]	Prior HCE and DCG model	Full model[#]	Incentive compatible model#	Subgroup size
Demographic information	No	Yes	Yes	Yes	Yes	Yes	
Prior LTC use	No	No	Yes	No	Yes	Yes	
Prior HCE and diagnosis information	No	No	No	Yes	Yes	Yes	
Activating support	3763	3556		3080			3559
Domestic care I	5006	-3232		-3524			46611
Domestic care II	9753	2147		1479			107181
Guidance	21909	11406		10617			9252
Personal care	16992	8293		7269			82116
Nursing	25799	17183		15303			24727
Residential home 1-90 days	22224	11978		10488		8207	10054
Residential home 91-180 days	34768	22153		20826			3903
Residential home 181-365 days	36137	23203		22153			7655
Residential home 366 days	31190	14777		14139			44726
Nursing home: somatic care 1-90 days	24292	16859		13829		13104	1042
Nursing home: somatic care 91-180 days	38721	30839		27338			389
Nursing home: somatic care 181-365 days	54610	46047		42608			376
Nursing home: somatic care 366 days	65995	56700		55923			1111
Nursing home: psychogeriatric care 1-90 days	51386	41756		40013		32614	755
Nursing home: psychogeriatric care 91-180 days	62188	52102		51072			385
Nursing home: psychogeriatric care 180-365 days	68104	57465		56406			517
Nursing home: psychogeriatric care 366 days	70693	58546		58887			2049
Nursing home: combined care 1-90 days	31009	23310		20888		21005	10546
Nursing home: combined care 91-180 days	51350	42482		39803			3853
Nursing home: combined care 180-365 days	71408	61349		59290			4359
Nursing home: combined care 366 days	85868	74655		74693			12439
Receiving home care on last day of 2004	10078	2459		1669		876	172291
Stay in LTC institution on last day of 2004	31213	21576		20881		3155	119785

[*]All reported predicted losses are significant at the $p < 0.05$ level; [#]Cells are empty if this variable is included in this risk adjustment model

*Table A3: predicted losses for subgroups based on HCE in 2000-2004**

	No risk adjustment	Demographic model	Prior LTC model	Prior HCE and DCG model#	Full model	Incentive compatible model#	Subgroup size
Demographic information	No	Yes	Yes	Yes	Yes	Yes	
Prior LTC use	No	No	Yes	No	Yes	Yes	
Prior HCE and diagnosis information	No	No	No	Yes	Yes	Yes	
In top 15% in prior 3 years: hospital + outpatient care	3410	1804	278				137618
In top 15% in prior 5 years: hospital + outpatient care	4031	2321	393				57597
Expenditures on medical equipment in prior year	4108	949	185			93	693144
Expenditures on medical equipment in prior 3 years	6524	1846	320				330813
Expenditures on medical equipment in prior 5 years	7259	2057	326				211030
Expenditures on transportation in prior year	7350	4271	582			916	201609
Expenditures on transportation in prior 3 years	9040	5753	353				43615
Expenditures on transportation in prior 5 years	8565	5347	374				22462
Expenditures on paramedical care in prior year	2409	1319	107			53	246233
In top 15% in prior 3 years: paramedical care	3635	2009	350				91743
In top 15% in prior 5 years: paramedical care	4371	2338	462				51218
In top 15% in prior year: pharmaceuticals	2832	572	130			28	701221
In top 15% in prior 3 years: pharmaceuticals	3215	678	140				482222
In top 15% in prior 5 years: pharmaceuticals	3448	742	147				381160
No prior HCE available	-445	25	17				1450876
DCG 1	2080	-272	-1474				22198
DCG 2	1921	466	117				27820
DCG 3	2400	349	117				14018
DCG 4	10172	6865	1145				11358
DCG 5	5109	2189	471				17981
DCG 6	3821	1346	276				10586
DCG 7	4595	3512	1215				7391
DCG 8	6187	3881	1033				7027
DCG 9	4787	3231	978				7459
DCG 10	7602	5840	3275				5365
DCG 11	7461	6220	3823				2939
DCG 12	5275	3536	1805				2684

*All reported predicted losses are significant at the $p < 0.05$ level; #Cells are empty if the variable is included in this risk adjustment model

Table A4: Predicted losses for subgroups based on diagnosis information from 2004 hospital admission data

Diagnosis in 2004	No risk adjustment	Demographic model	Prior LTC model[#]	Prior HCE and DCG model[#]	Full model	Incentive compatible model	Subgroup size
Demographic information	No	Yes	Yes	Yes	Yes	Yes	
Prior LTC use	No	No	Yes	No	Yes	Yes	
Prior HCE and diagnosis information	No	No	No	Yes	Yes	Yes	
10 diagnoses with the highest initial predicted loss							
Dementia	30423*	20777*	**9255***	20114*	9274*	11692*	821
Hip fracture	21225*	12029*	**-950***	9167*	-1229*	3205*	6433
Parkinson's disease	21093*	17566*	**6486***	13817*	5730*	7650*	510
Chronic ulcers of skin including decubitus	13421*	8676*	**-224**	4514*	-1897*	-1528*	873
Stroke	10840*	7288*	**1109***	2219*	327*	1271*	11998
Septicaemia	10620*	7322*	**1100***	4277*	592	2330*	679
Heart failure	10054*	3806*	858*	**745***	182	637*	8147
Pancreas cancer	9584*	7556*	6316*	**2605***	3525*	3469*	329
Osteoporosis	9334*	4590*	**1002***	2710*	662*	1660*	1598
Acute renal and urinary infections	8908*	5452*	**1252***	3420*	954*	1798*	3942
10 diagnosis with the highest initial predicted loss for which the incentive compatible model sufficiently reduced the predicted loss							
Heart failure	10054*	3806*	858*	**745***	182	637*	8147
Diseases of the blood and bloodforming organs	8139*	3713*	**1191***	1718*	494*	755*	6814
Diabetes mellitus including diabetic complications	6478*	4633*	**679***	2770*	133	231	4746
Stomach cancer	6148*	3900*	2673*	**-290**	368	553	664
Asthma and COPD	6128*	3721*	876*	**122**	-78	26	8035
Nephritis and nephropathy	5579*	3906*	**535***	767*	-189	128	2117
Esophagus cancer	5429*	3768*	2184*	**-1207**	-725	-343	426
Colorectal cancer	5416*	2438*	1095*	**-20***	235	464	3975
Other endocrine, nutritional and metabolic diseases	4504*	2829*	**761***	1440*	474*	849*	8510
Congenital anomalies of nervous system	4082	2558	**181**	546	-239	-229	50

* Significant at the $p < 0.05$ level; bold numbers indicate whether including prior LTC use or prior HCE and diagnosis information decreased the predicted loss most. Results for subgroups based on primary treatment information from 2004 and from earlier years and for subgroups based on primary diagnoses from earlier years available upon request

Table A5: predicted losses for subgroups based on the General Survey of Living Conditions (POLS) 2004

	No risk adjustment	Demographic model	Prior LTC model#	Prior HCE and DCG model#	Full model	Incentive compatible model	Subgroup size	Prevalence[a]
Demographic information		Yes	Yes	Yes	Yes	Yes		
Prior LTC use		No	Yes	No	Yes	Yes		
Prior HCE and diagnosis information		No	No	Yes	Yes	Yes		
From POLS survey (n = 9098)								
Self-rated health: bad	4591*	3167*	421	2443*	294	613	482	6.19
Long-term illness	1185*	380	32	186	-5	49	2927	37.6
Education: low	1314*	193	9	166	8	72	2616	33.95
Education: high	-834*	-245*	-20	-240*	-23	-86	1006	13.05
GALI problems[b]	2912*	1405*	134	986*	67	200	1327	17.05
Income: lowest quartile	661*	488*	95	474*	91	119	1936	
Income: highest quartile	-743*	-33	-1	6	3	-12	1945	
From POLS survey (n = 4619)								
Specialist visit in 2004	679*	157	11	-7	-19	-76	1584	43.77
Use of pharmaceuticals in prior 14 days	1039*	265	-158	25	-207	-294*	476	13.16
Physiotherapy	1042*	738*	-9	509	-36	-40	738	20.39
At least 1 chronic disease	270	111	63	101	60	10	5892	34.33
At least 2 chronic diseases	362*	223	86	218	85	46	5260	12.59
Blindness	638	-318	-189	-442	-209	-258	536	17.96
Deafness	1370*	-148	-124	-176	-134	-144	466	15.66
Diabetes	1509	-670	-60	-1153	-129	-320	121	4.05
Cancer	2621*	968	-26	588	-122	-319	181	6.07
Stroke	4349*	1404	-64	1161	-165	-302	108	3.61
Blood pressure	1801*	204	34	133	0	-53	405	13.5
Ciculatory system	2848*	466	-145	-146	-246	-477	77	2.57
Urinary incontinence	3131*	1283	683	911	631	688	166	5.53
Osteoarthritis	1649*	-126	-78	-266	-105	-124	445	14.82
Arthritis	2025*	451	-61	315	-98	-75	191	6.37
At least 1 ailment	198	-107	-66	-152	-77	-128*	1659	21.3
At least 2 ailments	752*	-49	-133	-167	-157	-194	917	11.77
PCS-12[c] score < 30	3746*	1834	-308	1326	-397	-219	208	7.62

Has difficulty to/cannot perform ≥ 1 ADL[d]	16221*	8918*	357	7887*	289	532	90	7.45
Cannot perform ≥ 1 ADL	23365*	13476*	1479	11906*	1409	1365	32	2.65
Has difficulty to/cannot perform ≥ 1 mobility task	9346*	3627*	212	2953*	156	37	195	16.28
Cannot perform ≥ 1 mobility task	15394*	7247*	464	6225*	388	656	69	5.76

*Predicted loss is not significant at the $p < 0.05$ level

[a]Prevalence among respondents for whom this characteristic is not missing

[b]Global Activity Limitation Indicator (van Oyen *et al.* 2006)

[c]Physical Component Scale (Ware *et al.* 1996)

[d]Activity of daily living: dressing; walking across a room; bathing; eating; getting in / out of bed; using the toilet (Katz and Akpom 1976)