Medical Care Research and Review
2020, Vol. 77(6) 584–595
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1077558719825982
journals.sagepub.com/home/mcr

Selection Incentives for Health Insurers in the Presence of Sophisticated Risk Adjustment

Richard C. van Kleef[1], Frank Eijkenaar[1], and René C. J. A. van Vliet[1]

Abstract

This article analyzes selection incentives for insurers in the Dutch basic health insurance market, which operates with community-rated premiums and sophisticated risk adjustment. Selection incentives result from the interplay of three market characteristics: possible actions by insurers, consumer response to these actions, and predictable variation in profitability of insurance contracts. After a qualitative analysis of the first two characteristics our primary objective is to identify the third. Using a combination of claims data (N = 16.8 million) and survey information (N = 387,195), we find substantial predictable variation in profitability. On average, people in good health are profitable, while those in poor health are unprofitable. We conclude that Dutch insurers indeed face selection incentives. A complete measure of selection incentives, however, captures the correlation between individual-level profitability and consumer response to insurer-actions. Obtaining insight in this correlation is an important direction for further research.

Keywords

health insurance, risk selection, risk equalization, risk adjustment

Introduction

The Netherlands has organized its basic health insurance scheme according to principles of regulated competition (van de Ven et al., 2013). Comparable schemes exist in Germany, Switzerland, and the United States. The model of regulated competition originates from the work by Alain Enthoven and combines competition among insurers with specific regulation to protect public objectives such as individual accessibility and affordability of coverage (Enthoven, 1978, 1988, 2012). In the Dutch scheme, competition is driven by a free consumer choice of health plan (resulting in competition among insurers) and freedom for insurers to decide where and by whom medical treatments are provided (resulting in competition among health care providers). Regulation includes a standardized benefits package in terms of medical services (such as primary care, cancer treatment, durable medical equipment, and pharmaceutical care), an insurance mandate, open enrollment, community rating per health plan, and risk adjustment (RA).

One of the main challenges in schemes with regulated competition is to avoid selection incentives for insurers. Risk selection by insurers may lead to efficiency problems and fairness issues (Glazer & McGuire, 2000; Rothschild & Stiglitz, 1976). In the Dutch context, efficiency problems can occur when insurers choose not to contract with providers who are relatively attractive to unprofitable consumers. Fairness issues can occur when differences in plan premiums do not only reflect variation in quality (e.g., in terms of provider network) and efficiency in production but also selection. Selection-driven premium variation can conflict with the regulator's concept of fairness in health care financing. Moreover, it may lead to inefficient sorting of consumers across plans (Akerlof, 1970; Einav & Finkelstein, 2011).

This article analyzes selection incentives for insurers in the Dutch basic health insurance. As will be explained in the next section, these incentives depend on the interplay of three market aspects: possible "actions" by insurers, variation in consumer response to these actions, and predictable variation in profitability of insurance contracts. After a

This article, submitted to *Medical Care Research* and Review on April 8, 2018, was revised and accepted for publication on December 30, 2018.

[1]Erasmus University Rotterdam, Rotterdam, Netherlands

Corresponding Author:
Richard C. van Kleef, Erasmus School of Health Policy & Management, Erasmus University Rotterdam, Postbus 1738, Rotterdam 3000 DR, Netherlands.
Email: vankleef@eshpm.eur.nl

qualitative analysis of the first two aspects, our primary objective is to identify the third. Using claims data covering the entire Dutch population (N = 16.8 million), we first replicate the risk-adjusted individual-level revenues that insurers receive for their enrollees and compare these with the medical spending of these enrollees. The gap between revenues and spending for an individual constitutes the insurer's profit (when revenues > spending) or loss (when revenues < spending) on that individual. In a second step, we combine these individual-level profits and losses with information from a health survey (N = 387,195 respondents of 19 years and older) to examine the extent to which variation in these profits and losses is predictable. More specifically, we calculate the mean profit/loss in year t for groups based on self-reported health measures from year $t - 1$.

The article proceeds as follows. The next section presents a conceptual framework for analyzing selection incentives. It discusses the role of the three aspects of selection incentives mentioned above and summarizes the contribution of our study. One of the key innovations of this paper is the combination of claims data with health survey information for 387,195 individuals. The content and size of this survey allows for the identification of profits and losses for health dimensions that are typically *not* taken into account in premiums and risk-adjusted payments. The section 'Predictable Variation in Profitability in the Dutch Health Insurance Market' reports on the methods and outcomes of the empirical analyses. Our key finding is that predictable variation in profitability is indeed present in the Dutch health insurance market. On average, groups in good health are profitable, while those in poor health are unprofitable. This brings us to the conclusion that some selection incentives for insurers are likely to exist in the Dutch health insurance market. We also conclude, however, that a complete measure of selection incentives should capture the correlation between individual-level profitability and consumer response to insurer-actions. In the Discussion section, we discuss some ideas for further research to obtain insight in this correlation.

Conceptual Framework

Selection Incentives: An Interplay of Three Market Characteristics

Newhouse (1996) defines risk selection as "Actions by consumers and insurers to exploit unpriced risk heterogeneity and break pooling arrangements." As argued by van Kleef, McGuire, Schut, and van de Ven (in press), the concept of "unpriced risk" from the consumers' perspective can differ from that of the insurers' perspective. In this article, we primarily focus on "unpriced risk" from the insurers' perspective, that is, the predictable variation in profitability of insurance contracts.

When it comes to selection incentives for insurers, Newhouse's definition implies three necessary conditions. First, insurers must be able to take "actions" that might lead to systematic sorting of different risk types into different plans. Second, there must be variation in how consumers respond to these insurer-actions. In a system where consumers respond *uniformly* to actions by insurers, no actions will lead to systematic sorting of different risk types into different plans. Third, there must be predictable variation in profitability of insurance contracts. In a system without such variation, systematic sorting of profitable and unprofitable consumers into different plans is absent by definition. Below, we take a closer look at each of these three aspects and discuss the extent to which they are (likely to be) present in the Dutch context.

Possible Actions by Insurers

By "insurer-actions," we mean all possible actions on the side of insurers that can lead to sorting of profitable and unprofitable people into different plans. At least six types of actions are possible in the Dutch basic health insurance market. First, insurers have flexibility with respect to provider network design. While the nature and content of covered benefits are determined by the government (e.g., primary care, hospital care, and prescribed drugs)—insurers are free to decide where, by whom and under which conditions treatments are to be provided. Second, insurers can offer cost sharing options. The Dutch basic health insurance includes a mandatory deductible of 385 euro per adult per year in 2016. On top of this deductible, insurers are allowed to offer voluntary deductibles of 100, 200, 300, 400 and/or 500 euros per adult per year. In addition, insurers can charge copayments for out-of-network expenditures. Third, insurers have substantial flexibility when it comes to utilization management and provider remuneration. Examples of utilization management are that insurers can actively assist patients in choosing a provider or require a second opinion for certain treatments before consumers get access to reimbursement. A more indirect way of utilization management concerns the contractual arrangements between insurers and providers in terms of quality requirements and remuneration methods. Fourth, insurers have freedom in terms of customer service. Differences in customer service among plans can occur in terms of options to contact the office—for example, in person, by telephone or exclusively via the Internet—and query–response time. Fifth, insurers have much flexibility in terms of sales and marketing. For example, advertisement can be targeted at particular groups of consumers. Insurers can also provide special privileges to people who enroll via a so-called "group arrangement." Such arrangements can be organized by any legal entity, for example, employers, shops, sports clubs, patient organizations, and private initiatives. Sixth, insurers can offer supplementary insurance. Though basic health insurance and supplementary health insurance must be contractually separated by law, consumers tend to perceive these as a single product (Duijmelinck & van de Ven, 2014). When consumers' preferences regarding supplementary insurance are correlated with their profitability in basic insurance, variation in characteristics of the first can be a mechanism for risk selection regarding the latter.

Variation in Consumer Response to Insurer-Actions

When it comes to selection incentives for insurers, the role of consumer response has been clearly described in theoretical and empirical papers on "service-level selection," that is, the phenomenon that insurers design their plans in a way to attract profitable consumers and/or deter unprofitable ones. When insurers decide to reduce investments in the availability or quality of a certain service, consumer response will most likely be some function of predicted spending for that service (Ellis & McGuire, 2007). Based on an economic model of profit maximization developed by Frank, Glazer, and McGuire (2000), Ellis and McGuire (2007) construct and apply measures to identify incentives for service-level selection by making assumptions about consumer response. In recent publications, McGuire, Newhouse, Normand, Shi, and Zuvekas (2014) and Ellis, Martins, and Zhu (2017) replace some of these assumptions with empirical estimates of demand elasticities. A crucial insight from this line of research is that incentives for service-level selection follow from the correlation between individual-level profitability and consumer response to insurer-actions. In fact, this is not only true for service-level selection but also for any other action on the side of insurers.

Consumer response to insurer-actions can depend on various factors, such as consumers' prediction of spending for different types of medical services (Ellis & McGuire, 2007), attitude toward risk, transaction costs, price sensitivity, quality preferences, and knowledge of the health care system. Moreover, these factors are likely to be interdependent and influenced by a variety of underlying characteristics, such as education and income. Though precise figures on the relationship between insurer-actions and consumer response are mostly absent, some general patterns can be observed in the Dutch basic health insurance market.

By analyzing data of nearly the entire Dutch population, Duijmelinck and van de Ven (2016) find that in 2009 consumers in the age group 25 to 44 years had a 10 times higher switching rate than the group of 75 years and older. Moreover, they find that switching rates decrease with predicted spending. More specifically, healthy consumers switch twice as much as the nonhealthy, though these differences become much smaller after adjusting for age. The authors explain these findings by higher perceived switching costs by elderly consumers than by younger consumers.

More specific evidence of consumer response is provided by studies investigating the correlation between predicted spending and deductible choice. A recent study found that in the Dutch health insurance market, healthy individuals more often opt for a voluntary deductible with a community-rated premium rebate than people in poor health (Centraal Planbureau, 2016). Based on similar data as used in this article, the authors find that average medical spending in 2013 is considerably lower for people with the highest voluntary deductible than for those without a voluntary deductible: 450 versus 2.350 euros per person per year.

In a recent study on consumer preferences, Bes, Curfs, Groenewegen, and de Jong (2017) find that those who are willing to choose a restrictive health plan in return for a lower premium are on average younger and healthier than those who prefer a nonrestrictive plan. Though these findings are only based on stated preferences, they do indicate variation in consumer response.

The aforementioned studies do not directly indicate how consumers respond to specific insurer-actions in the Dutch context. Despite the lack of empirical research, there is some anecdotal evidence. An exemplary case comes from a health insurer that offered a supplementary insurance plan that was particularly attractive to pregnant women expecting to deliver a baby in the contract year. Indeed, this insurer attracted relatively many pregnant women. Since the Dutch RA model lacks risk adjustors that explicitly indicate pregnancy and compensate for the associated costs, this insurer was confronted with substantial losses. This led the insurer to discontinue the plan and also discouraged other insurers to start offering such products.

Predictable Variation in Profitability of Insurance Contracts

When variation in expected spending across insurance contracts is not reflected in the revenues insurers receive for these contracts, some contracts will on expectation be profitable (when revenues > expected spending), while others will be unprofitable (when revenues < expected spending). As in most health insurance markets based on regulated competition, premiums for the Dutch basic health insurance must be community rated per health plan. In other words, people choosing the same plan pay the same premium, regardless of their expected spending. Without further measures, this would lead to substantial predictable variation in profitability, since expected spending of the elderly and chronically ill far exceed that of the young and healthy. Consequently, the elderly and chronically ill would be very unprofitable to insurers, while the opposite holds for the young and healthy. To compensate insurers for these predictable profits and losses, the Dutch basic health insurance scheme includes a system of RA (also known as risk equalization). As will be described in more detail in later on, the Dutch RA system provides insurers with risk-adjusted payments based on age, gender, a broad range of health indicators, and a series of socioeconomic variables. Our empirical analysis below aims at identifying predictable variation in profitability that remains *after* RA.

New Contributions

A main challenge that comes with identifying predictable variation in profitability is to obtain information on risk factors that is not yet taken into account in the RA model. In this article, we overcome this challenge by combining administrative data on insurance claims and risk characteristics with information from a health survey. More specifically, we replicate the Dutch RA model of 2016 using administrative data on all individuals with a health plan for the basic health insurance package in 2013 (N = 16.8 million). This allows for calculating the revenues insurers receive for their enrollees. We then merge individual-level revenues and actual claims with health survey information from 2012 (N = 387,195 of 19 years and older), which makes it possible to determine payment fit for different sets of mutually exclusive groups based on self-reported health (both physical and mental) and lifestyle. A key novelty of this article is that these partitions of the population are typically impossible to create using claims data alone. Moreover, the large number of respondents allows for calculating payment fit much more robustly than in previous studies. For example, van Kleef, van Vliet, and van de Ven (2013) used a health survey with only 15,000 respondents to analyze payment fit. Another innovation is that we identify predictable variation in profitability conditional on one of the most sophisticated RA models in the world. If the Dutch RA model does not completely compensate for variation in predictable spending, the same is likely to be true for other—less sophisticated—RA models used in other individual insurance markets (ceteris paribus).

Predictable Variation in Profitability in the Dutch Health Insurance Market

Our empirical analysis identifies predictable variation in profitability using a three-step procedure. First, we estimate the RA model of 2016 using the administrative data and calculate individual-level predicted spending. Second, we merge predicted and actual spending with the health survey information. And third, we calculate the mean difference between actual and predicted spending for different sets of mutually exclusive groups based on self-reported health. Below, we describe these steps in more detail and present the main results.

Estimating the Risk Adjustment Model of 2016

Since premiums for the Dutch basic health insurance must be community rated per health plan, fit between revenues and spending can roughly be indicated by the residual spending from the RA model. For simplicity, we refrain from the loading fee and other types of revenue. In 2016, the Dutch RA system consisted of four different models, one for each of the following categories: somatic care, short-term mental health care (i.e., mental treatments for people who are not institutionalized), long-term mental health care (i.e., mental treatments for people who are institutionalized), and out-of-pocket payments due to the mandatory deductible. Each of these four models leads to an individual-level prediction of the relevant spending type. *Total* predicted spending under the basic health insurance $\hat{y}$ for individual i is calculated as:

$$\hat{y}_{i,total} = \hat{y}_{i,somatic} + \hat{y}_{i,mental\_short} + \hat{y}_{i,mental\_long} - \hat{y}_{i,oop} \quad (1)$$

with the four components on the right-hand side referring to i's predicted spending for somatic care, short-term mental care, long-term mental care, and out-of-pocket payments due to the mandatory deductible, respectively. Similarly, *total* actual spending y for i equals:

$$y_{i,total} = y_{i,somatic} + y_{i,mental\_short} + y_{i,mental\_long} - y_{i,oop} \quad (2)$$

In 2016, the RA model for somatic care included 162 risk classes based on the following characteristics: age interacted with gender, region, source of income interacted with age, pharmacy-based cost groups (PCGs), diagnosis-based cost groups (DCGs), socioeconomic status interacted with age, multiple-year high cost (MYHC), durable medical equipment cost groups (DME), and groups based on prior-year spending for specific services such as home care. For a detailed description of these risk adjustors, see van Kleef, Eijkenaar, van Vliet, and van de Ven (2018).

In 2016, the RA model for short-term mental health care included 95 risk classes based on the following characteristics: age interacted with gender, region, source of income interacted with age, PCGs for mental diseases, DCGs for mental diseases, socioeconomic status interacted with age, household size interacted with age, and MYHC for mental care. The RA model for long-term mental care mimicked the model for short-term mental health care with one additional risk adjustor: spending for inpatient mental health care in the previous year (van Kleef et al., 2018). Both models for mental health care solely apply to people of 18 years and older. For people younger than the age of 18 years, mental health care is financed by a public program.

The RA model for out-of-pocket spending is applied to correct RA payments for predictable variation in out-of-pocket spending under the mandatory deductible. This is necessary because the RA models for somatic care and mental care lead to predictions of total spending including the out-of-pocket payments rather than spending covered by the plan. In 2016, the mandatory deductible was 385 euro per adult per year. The deductible applies to all health care services covered by the benefits package—both somatic and mental care—except for primary care provided by general practitioners, maternity care, obstetrics, and home care. In terms of risk adjustors, the model relies on the following information: age interacted with gender, source of income, and region. The model only applies to people of 18 years and older (since those younger than 18 years are exempted from the deductible) and to those without a PCG, DCG, MYHC, and DME (all from the somatic RA model). For individuals with a PCG, DCG, MYHC, and/or DME—whose spending generally exceeds the deductible—the predicted out-of-pocket payments equal the average out-of-pocket spending in this group, which nearly equals the deductible amount.

In all four RA models risk classes take the form of dummy variables. Risk-adjustor coefficients are derived by an individual-level regression of spending in 2013 on risk characteristics (i.e., the dummy variables) from 2013 (age, gender, source of income, socioeconomic status, and region), from 2012 (PCGs, DCGs, DME, and groups based on prior spending for specific services), or before (MYHC, which in the model for somatic care is based on spending levels in the years 2010-2012 and in the model for mental care on spending levels in the years 2008-2012). Data on medical spending and risk adjustors cover the entire Dutch population with a health plan in 2013. Prior to estimation, some modifications were applied to make the lagged data representative for 2016, such as corrections for changes in the benefits package between 2013 and 2016. For both the somatic model and the model for out-of-pocket payments under the mandatory deductible, risk-adjustor coefficients are derived by an ordinary least-square regression. For the mental care models, coefficients are derived by quadratic programming with the restriction that predictions must be positive. For all models, the regression is based on annualized medical spending weighted by the fraction of the year an individual was enrolled in 2013. This fraction can be smaller than 1.0 due to birth, death, migration, and within-year switching of plans which occasionally occurs, for example, when children turn 18 years and obtain the right to choose their own plan.

For the purpose of this study, we obtained permission to use the administrative data from the period 2008-2013 to replicate the RA models of 2016. This data set includes individual-level information on medical spending and risk characteristics for the entire population under the Dutch basic health insurance of 2013 (N = 16.8 million). The information in this data set comes from various administrative sources, including insurers, the tax collector, and the registration service for social benefits. After replicating the RA models, we were able to calculate individual-level predicted spending according to Equation (1). Since the data set includes information on actual spending as well, we were able to calculate the insurer's financial result for individual i as

$$financial\ result_i = \hat{y}_{i,total} - y_{i,total} \tag{3}$$

Given that premiums in the Dutch health insurance market are community rated per health plan, overall payment system fit can be approximated by comparing the variation in financial result of Equation (3) with the variation in actual spending (Layton, Ellis, McGuire, & van Kleef, 2017). In a simulation of plan revenues and spending, we found that the four RA models together compensate for 29.8% of the *squared* deviations between individual-level spending and mean spending. For the *absolute* deviations between individual-level spending and mean spending, this figure equals 29.7%.

Merging Individual-Level Results With Health Survey Information

In a next step, we combined the individual-level spending and financial results with health survey information from 2012. For the purpose of this research, we were able to merge the different data sets on the basis of unique, individual-level identification codes which for privacy reasons were anonymized by a trusted third party. For a total of 387,195 respondents, the survey data contain information on self-reported general health (both physical and mental), chronic conditions, and lifestyle (e.g., smoking and alcohol intake). For 99.2% of the individuals in the survey sample, we found a successful match with the administrative data. Reasons for an unsuccessful match are death and migration in 2012 and nonenrollment in the Dutch basic health insurance of 2013, for example, small groups of defaulters and military servants.

The survey sample is in fact a combination of three surveys: the adult monitor (19-64 years), the elderly monitor (65 years and older), and the health monitor by Statistic Netherlands (19+ years). The latter consists of two parts: a first set of questions sent to all respondents and a second set of questions sent only to those who were willing to take part in the follow-up. For three reasons, the composition of the survey sample differs from that of the population. First, the survey was not sent to people living in an institution for long-term care. Second, the total sample only includes people of 19 years or older. Third, respondents in the remaining population were not selected randomly, which resulted in an overrepresentation of some groups (e.g., the elderly) and an underrepresentation of others. To correct for possible selection, Statistics Netherlands reweighted the sample on the basis of age, gender, marital status, degree of urbanization, household size, ethnicity, income, and region. In addition, we applied an iterative proportional fitting procedure to rebalance the survey sample in such a way that the weighted sample frequencies of risk adjustor variables equal those in the population (Battaglia, Hoaglin, & Frankel, 2009). In addition to the risk adjustor variables, the rebalancing procedure also took into account a grouping of the population into quantiles of medical spending and a partition based on a proxy for yes/no deceased in 2013. The following example illustrates how the procedure works. Assume the survey is to be rebalanced on the basis of age (i.e., 14 groups in our empirical analysis) and gender only. In the first step, the weight of each case in the sample is multiplied by the ratio of the population frequency to the weighted sample frequency of the relevant age group. This step results in reweighted sample frequencies for all age groups that agree with population frequencies. In the next step, the new weight of each

Table 1. Descriptive Statistics of Spending and Characteristics in 2013: (Rebalanced) Survey Sample Versus Population (All 19+ Years).

	Survey sample (19+ years)	Rebalanced survey sample (19+ years)	Population (19+ years)
Number of insured	384,004	384,004	12,926,184
Number of insured years	381,283	12,774,890	12,774,877
Mean actual spending, €	3,116*	2,478	2,493
Mean predicted spending according to RA model 2016, €	3,171*	2,494	2,493
Mean financial result (i.e., predicted spending—actual spending), €	55*	16	0
Men, %			
19-34 Years	5.6*	11.8	11.8
35-44 Years	4.8*	8.7	8.7
45-54 Years	6.6*	9.8	9.8
55-64 Years	7.6*	8.4	8.4
65 Years or older	20.7*	10.1	10.1
Women, %			
19-34 Years	8.0*	11.8	11.8
35-44 Years	6.5*	8.8	8.8
45-54 Years	8.3*	9.8	9.8
55-64 Years	8.5	8.4	8.4
65 Years or older	23.4*	12.4	12.4

Note. RA = risk adjustment. Means of actual and predicted spending are presented per "insured year." Frequencies of age/gender groups are calculated as a percentage of total insured years in the population and survey sample, respectively.
*Statistically significantly different from population mean ($p < .05$).

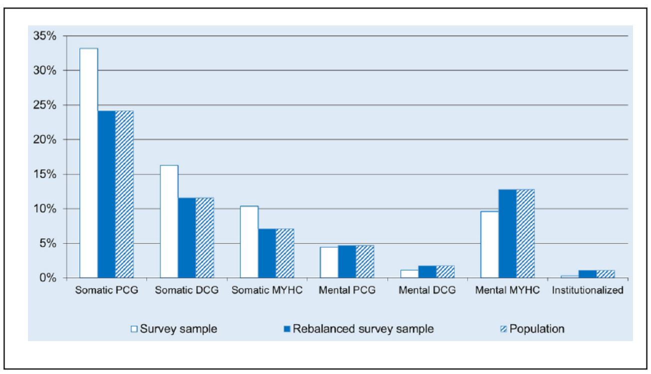

Figure 1. Relative frequency of specific indicators: Survey sample versus population (19+ years).
Note. PCG = pharmacy-based cost group; DCG = diagnosis-based cost group; MYHC = multiple-year high cost; Institutionalized = living in an institution for long-term care.

case is multiplied by the ratio of the population frequency to the reweighted sample frequency of the corresponding gender. However, now the reweighted sample frequencies for the *age* groups do not agree with the corresponding population frequencies anymore, but they are closer then at the start. This process is repeated until the weighted frequencies for both age groups and gender in the sample agree with the population frequencies.

Table 1 compares spending and characteristics in the survey sample with those in the total population of 19 years and older. In the sample, the mean actual spending in 2013 is substantially higher than in the population: 3,116 versus 2,493 euros per person. The same is true for the mean predicted spending according to the RA model: 3,171 versus 2,493 euros per person. These differences are due to the overrepresentation of elderly people in the sample. In the rebalanced sample, frequencies of age groups exactly equal those in the population (see bottom half of Table 1). Actual and predicted spending in the rebalanced sample nearly equal those in the population; remaining differences are no longer statistically significant.

Figure 1 compares the survey sample and the total population of 19+ years in terms of the frequencies of several indicators included in the RA model. In terms of indicators related to somatic care the unbalanced survey sample is overrepresented by people with morbidity. The opposite holds for indicators related to mental care and people living in an institution for long-term care. The underrepresentation of institutionalized people (in 2013) is due to the fact that institutionalized people (in 2012) were not selected for the health survey. Nevertheless, we find that the survey does contain respondents who were institutionalized in 2013. Apparently, these people moved to an institution after they completed the survey in 2012. In the rebalanced sample, the frequencies for the indicators in Figure 1 exactly equal those in the total population of 19+ years.

While Figure 1 presents the relative frequencies of risk-adjustor groups in the RA model of 2016, Figure 2 presents the mean actual spending in these groups in 2013. Except for the group of people who were institutionalized in 2013, the mean spending in the survey sample is relatively close to that in the total population of 19+ years. After rebalancing the

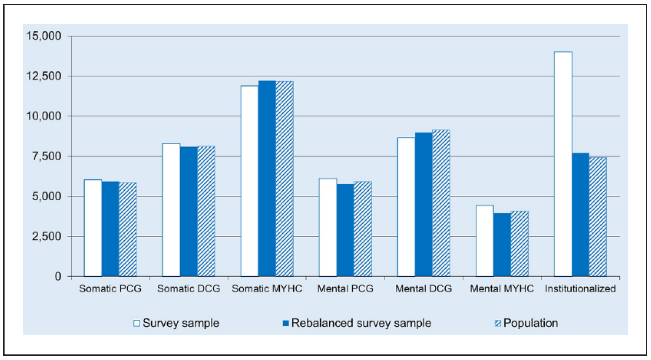

Figure 2. Mean actual spending per indicator in euros: Survey sample versus population (19+ years).
Note. PCG = pharmacy-based cost group; DCG = diagnosis-based cost group; MYHC = multiple-year high cost; Institutionalized = living in an institution for long-term care.

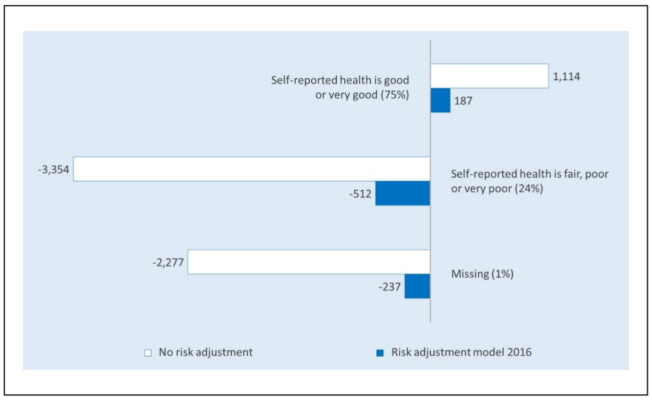

Figure 3. Mean financial result in euros per person per subgroup per year (19+ years).
Note. Percentages refer to population frequencies in the rebalanced survey sample. Mean result per person per year is determined as mean predicted spending per person per year minus mean actual spending per person per year.

sample, remaining discrepancies are almost eliminated, even for those who were institutionalized.

From the results above, we conclude that the mean (predicted) spending in the survey sample is substantially higher than in the total population of 19 years and older, the reason being the overrepresentation of elderly people. Our sample-rebalancing procedure nearly eliminates these discrepancies, both at the sample level and at the level of specific risk-adjustor groups.

Calculating Payment Fit for Various Sets of Mutually Exclusive Groups

As a final step, using Equation (3), we calculated the mean of the financial result (henceforth: mean result) for various sets of mutually exclusive groups identifiable in the survey sample. The groupings are based on four types of indicators: (a)

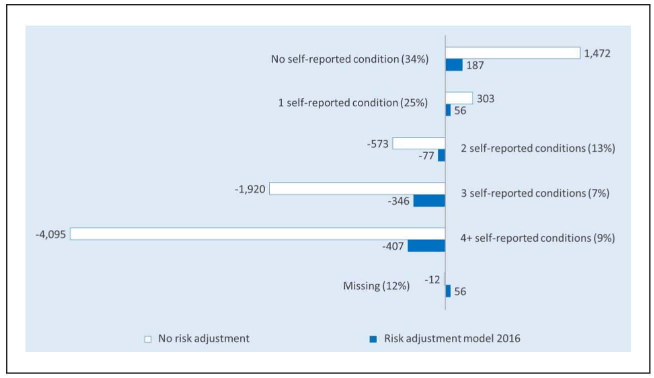

Figure 4. Mean financial result in euros per person per subgroup per year (19+ years).
Note. Percentages refer to population frequencies in the rebalanced survey sample. Mean result per person per year is determined as mean predicted spending per person per year minus mean actual spending per person per year.

general self-reported health, (b) the number of self-reported conditions, (c) the risk of incurring an anxiety disorder or depression, and (d) the level of physical activity. For each set of mutually exclusive groups, we present the estimated population frequency, as well as the mean result. The mean result is calculated both *pre*- and *post*-RA. The former indicates the predictable variation in profitability under community-rated premiums alone, while the latter indicates the predictable variation in profitability that remains after application of the Dutch RA model 2016. Note that survey information from year $t - 1$ can never fully predict residual spending in year t due to random variation in spending, the subjective nature of self-reported health and changes in (self-reported) health between year $t - 1$ and year t. The goal of our analysis is to identify *predictable* differences.

Figure 3 shows the mean results for people who report a fair, poor, or very poor health status (24% of the sample) and those who report a good or very good health status (75% of the sample). Without RA but with community rating, the difference in mean result between these groups would be 4,468 euros per person per year, indicating considerable predictable variation in profitability. RA reduces this difference to 699 euros, implying that the RA model of 2016 substantially reduces predictable profits and losses, but not completely. For 1% of the survey sample data on self-reported health are missing. For consistency, results for the group with missing values are reported in each of the Figures 3 to 6. Consequently, the sum product of relative frequencies and results is similar across these figures and consistent with the results in Table 1.

Figure 4 presents the mean result for a set of mutually exclusive groups based on the number of self-reported conditions. An overview of the mean result for each of the underlying conditions can be found in the appendix. The differences in mean result without RA among the groups in Figure 4 indicate considerable predictable variation in profitability. Also for this grouping, RA substantially reduces predictable

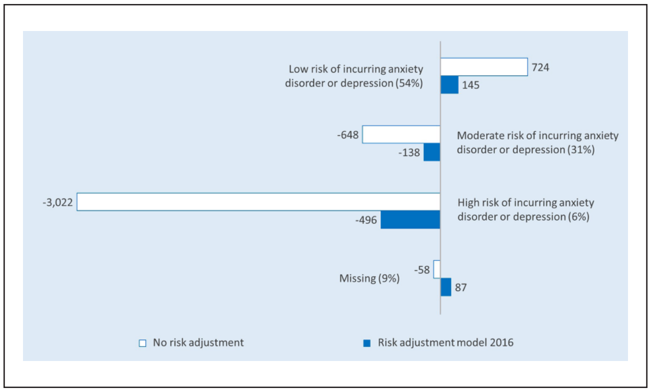

Figure 5. Mean financial result in euros per person per subgroup per year (19+ years).
Note. Percentages refer to population frequencies in the rebalanced survey sample. Mean result per person per year is determined as mean predicted spending per person per year minus mean actual spending per person per year.

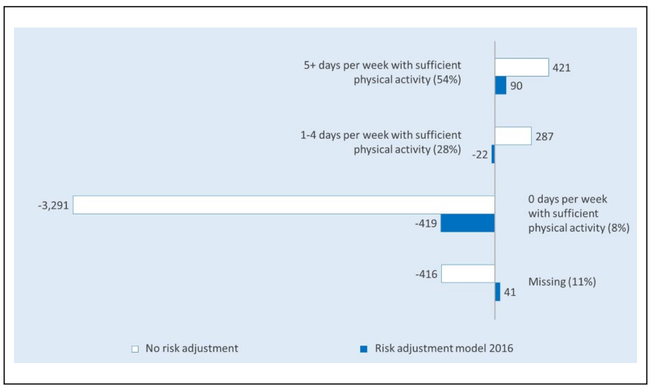

Figure 6. Mean financial result in euros per person per subgroup per year (19+ years).
Note. Percentages refer to population frequencies in the rebalanced survey sample. Mean result per person per year is determined as mean predicted spending per person per year minus mean actual spending per person per year.

profits and losses, but not completely. For example, the difference in mean result between the groups "no self-reported condition" and "four or more self-reported conditions" reduces from 5,567 to 594 euros per person per year.

Figure 5 shows the mean result for a set of mutually exclusive groups related to mental health or, more specifically, the risk of incurring an anxiety disorder or depression. This indicator is constructed from respondents' answers to a series of questions about their mood in the past weeks. Again, RA substantially reduces predictable profits and losses, but not completely. The difference in mean result between the groups with a low, respectively, high risk of incurring an anxiety disorder or depression reduces from 3,746 to 641 euros per person per year.

Figure 6 presents the mean result for a set of mutually exclusive groups based on the number of days per week with

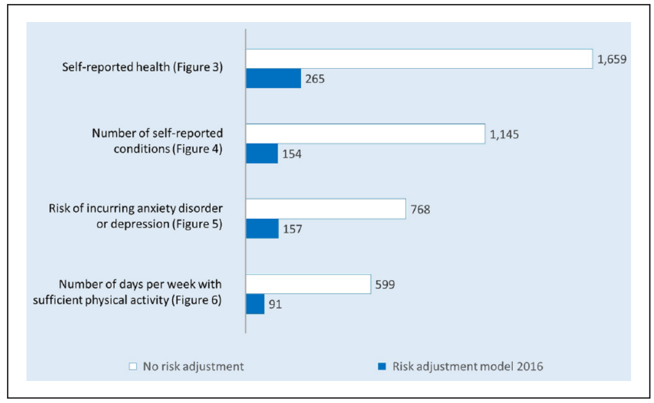

Figure 7. Mean absolute financial result in euros per person per partition per year (19+ years).
Note. Mean absolute financial result for a partition is calculated as the weighted sum of the absolute values of the mean result for subgroups in that partition (including a separate subgroup for individuals with missing values). For example, the mean absolute result without risk adjustment for the partition based on self-reported health (Figure 3) is calculated as 75.14% × 1,114 + 23.78% × 3,354 + 1.07% × 2,277 = 1,659.

"sufficient" physical activity. Again, RA reduces predictable profits and losses, but not completely. An interesting observation is that RA shrinks the mean results, but never alters the ordering of the groups. This is also true for the groupings in Figures 3 to 5.

Figure 7 summarizes the predictable variation in profitability *pre*- and *post*-RA for each of the sets of mutually exclusive groups presented in Figures 3 to 6. The bars show the weighted mean absolute result (WMAR) over the relevant set of groups. For example, the WMAR without RA for the grouping based on self-reported health (for which mean results are presented in Figure 3) is calculated as $0.7514 \times 1{,}114 + 0.2378 \times 3{,}354 + 0.0107 \times 2{,}277 = 1{,}659$. The difference in WMAR with and without RA approximates the extent to which RA reduces predictable variation in profitability. This reduction is 84% for the grouping based on self-reported health, 87% for the number of self-reported conditions, 80% for the risk of incurring anxiety disorder or depression and 85% for the number of days per week with sufficient activity.

We conclude that without RA, community rating in the Dutch health insurance market would lead to considerable predictable variation in profitability. RA substantially reduces predictable profits and losses, but not completely. Regarding the four sets of mutually exclusive groups presented in this section, RA on average reduces the predictable variation in profitability by 84% (calculated as the mean reduction in WMAR for the four sets of groups in Figure 7).

Discussion

In this article, we looked at three determinants of selection incentives for insurers in a health insurance market with community rating per health plan and RA: possible insurer-actions, variation in consumer response to these actions, and predictable variation in profitability of insurance contracts. Based on existing literature and original empirical research, we conclude that all three determinants are (likely to be) present in the Dutch basic health insurance market. Our qualitative assessment of that scheme shows there are many possible actions on the side of insurers that may lead to selection. In addition, based on empirical literature about health plan choice and switching, we conclude that variation in consumer response to insurer-actions is likely to exist as well. Finally, our new empirical results show that predictable variation in profitability is also present: After RA selected groups in good health tend to be profitable, while selected groups in poor health tend to be unprofitable.

The Dutch RA model is one of the most sophisticated RA models in the world. Our finding that predictable variation in profitability exists in this setting could mean that such variation is also present in other markets with a comparable benefits package and stringent premium regulation. Note, however, that the extent to which predictable profits and losses can actually result in risk selection depends on the possible insurer-actions and consumer response to these actions. For example, in markets where insurers have substantial flexibility regarding health plan design (such as the U.S. marketplaces) predictable variation in profitability might generate bigger or more selection problems than in markets where insurers have little flexibility regarding health plan design (such as the basic health insurance market in Germany and that in Belgium).

Though our findings indicate that all three ingredients for selection incentives are present in the Dutch health insurance market, a complete measure of selection incentives for insurers requires the ability to connect these aspects. More specifically, selection incentives depend on the correlation between individual-level profitability and consumer response to specific insurer-actions. Identifying such correlation requires individual-level data on plan revenues, medical spending and expected consumer response to specific actions. In the Dutch context, information on plan revenues and spending is readily available. Individual-level information on expected consumer response to particular actions, however, is not yet available. We see two options for obtaining such information. First, researchers could exploit health plan enrollment data, which has been collected since the introduction of the basic health insurance in 2006. By combining such information with observed insurer-actions (e.g., in terms of marketing and plan design), it may be possible to link health plan choice to particular actions. A disadvantage of this approach is that in the likely case of multiple actions, it might not be possible to isolate the effect of a particular action. Moreover, observed past behavior is not necessarily a good predictor for future behavior. Both problems might partly be overcome by mapping stated preferences (the second approach), for example, via surveys or discrete choice experiments (DCE's). Compared with the first approach, surveys and DCE's provide better opportunities for estimating the partial effect of certain actions. More specifically, the regulator could consider conducting an annual survey or DCE in which consumers are provided with hypothetical questions regarding provider network and other plan features. By merging survey outcomes with individual-level spending, the correlation between individual-level profitability and preferences could then be identified.

Further reduction of selection incentives is crucial for the functioning of health insurance markets. In general, there are two approaches to realize this. First, the regulator can limit the set of insurers-actions, for example, by further standardizing health plans in terms of coverage and provider network. A disadvantage, however, is that this approach not only reduces selection problems but also limits insurers' instruments for efficiency as well as consumer choice. Therefore, a better approach is to mitigate predictable variation in profitability by modifying the payment system, for example, by extending RA with new risk adjustor variables, supplementing RA with risk sharing (e.g., excess–loss compensations) and/or providing insurers with some flexibility to risk rate their premiums (e.g., a premium bandwidth).

Acknowledgments

The authors gratefully acknowledge the valuable comments on earlier versions of this article by two anonymous reviewers, Thomas McGuire, Mieke Reuser, Suzanne van Veen, Wynand van de Ven, Anja Withagen-Koster, and the participants of the Risk Adjustment Network meeting 2017. The authors are also grateful to the Dutch Ministry of Health, Welfare and Sports, the Dutch Association of Health Insurers, and the Dutch local health authorities for providing the claims data and health survey information for this study.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Richard C. van Kleef https://orcid.org/0000-0002-8971-6341

References

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*, 488-500.

Battaglia, M. P., Hoaglin, D. C., & Frankel, M. R. (2009). Practical considerations in raking survey data. *Survey Practice*, *2*(5), 1-14. Retrieved from https://www.surveypractice.org/article/2953-practical-considerations-in-raking-survey-data

Bes, R., Curfs, E., Groenewegen, P., & de Jong, J. (2017). Health plan choice in the Netherlands: Restrictive health plans preferred by young and healthy individuals. *Health Economics, Policy and Law*, *12*, 345-362.

Centraal Planbureau. (2016). *Keuzegedrag verzekerden en risicosolidariteit bij vrijwillig eigen risico* [Consumer behavior and risk solidarity regarding a voluntary deductible]. The Hague, Netherlands: Author.

Duijmelinck, D. M. I. D., & van de Ven, W. P. M. M. (2014). Choice of insurer for basic health insurance restricted by supplementary insurance. *European Journal of Health Economics*, *15*, 737-746.

Duijmelinck, D. M. I. D., & van de Ven, W. P. M. M. (2016). Switching rates in health insurance markets decrease with age: Empirical evidence and policy implications from the Netherlands. *Health Economics, Policy and Law*, *11*, 141-159.

Einav, L., & Finkelstein, A. (2011). Selection in insurance markets: Theory and empirics in pictures. *Journal of Economic Perspectives*, *25*, 115-138.

Ellis, R. P., Martins, B. D., & Zhu, W. (2017). Health care demand elasticities by type of service. *Journal of Health Economics*, *55*, 232-243.

Ellis, R. P., & McGuire, T. G. (2007). Predictability and predictiveness in health care spending. *Journal of Health Economics*, *26*, 25-48.

Enthoven, A. C. (1978). Consumer-choice health plan: A national-health-insurance proposal based on regulated competition in the private sector. *New England Journal of Medicine*, *298*, 709-720.

Enthoven, A. C. (1988). Managed competition of alternative delivery systems. *Journal of Health Politics, Policy and Law*, *13*, 305-321.

Enthoven, A. C. (2012). *Health care, the market and consumer choice*. Cheltenham, England: Edward Elgar.

Frank, R., Glazer, J., & McGuire, T. G. (2000). Measuring adverse selection in managed health care. *Journal of Health Economics*, *19*, 829-854.

Glazer, J., & McGuire, T. G. (2000). Optimal risk adjustment of health insurance premiums: An application to managed care. *American Economic Review*, *90*, 1055-1071.

Layton, T. J., Ellis, R. P., McGuire, T. G., & van Kleef, R. C. (2017). Measuring efficiency of health plan payment systems in managed competition health insurance markets. *Journal of Health Economics*, *56*, 237-255.

McGuire, T. G., Newhouse, J. P., Normand, S.-L., Shi, J., & Zuvekas, S. (2014). Assessing incentives for service-level selection in private health insurance exchanges. *Journal of Health Economics*, *35*, 47-63.

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

Rothschild, M., & Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, *90*, 629-649.

van de Ven, W. P. M. M., Beck, K., Buchner, F., Schokkaert, E., Schut, F. T., Shmueli, A., & Wasem, J. (2013). Preconditions for efficiency and affordability in competitive healthcare markets: Are they fulfilled in Belgium, Germany, Israel, the Netherlands and Switzerland? *Health Policy*, *109*, 226-245.

van Kleef, R. C., Eijkenaar, F., van Vliet, R. C. J. A., & van de Ven, W. P. M. M. (2018). Health plan payment in the Netherlands. In T. G. McGuire & R. C. van Kleef (Eds.), *Risk adjustment, risk sharing and premium regulation in health insurance markets: Theory and practice* (pp. 397-429). Amsterdam, Netherlands: Elsevier.

van Kleef, R. C., McGuire, T. G., Schut, F. T., & van de Ven, W. P. M. M. (in press). Strategies to counteract risk selection in social health insurance markets.

van Kleef, R. C., van Vliet, R. C. J. A., & van de Ven, W. P. M. M. (2013). Risk equalization in the Netherlands: An empirical evaluation. *Expert Review of Pharmacoeconomics & Outcomes Research*, *13*, 829-839.

Appendix

Table A1. Relative Population Frequency, Mean (Predicted) Spending, and Mean Financial Result in Euros Per Person Per Year (19+ Years) for Various Sets of Mutually Exclusive Groups Based on Self-Reported Conditions.

Self-reported condition in year 2012	Population frequency, %	Mean spending in 2013, €	Mean predicted spending for 2013 according to the RA model 2016, €	Mean financial result in 2013 under the RA model 2016, €
Diabetes (ever)				
Yes	5.8	6,626	6,499	−127*
No	88.1	2,160	2,187	27*
Missing	6.1	3,121	3,120	−1
Stroke (ever)				
Yes	2.9	7,966	7,060	−906**
No	91.5	2,277	2,322	45**
Missing	5.7	2,934	2,949	15
Acute myocardial infarction (ever)				
Yes	3.0	7,589	7,210	−379**
No	91.3	2,278	2,310	32**
Missing	5.7	3,008	2,983	−25

(continued)

Table A1. (continued)

Self-reported condition in year 2012	Population frequency, %	Mean spending in 2013, €	Mean predicted spending for 2013 according to the RA model 2016, €	Mean financial result in 2013 under the RA model 2016, €
Cancer (ever)				
Yes	6.5	6,454	6,028	−426**
No	88.0	2,165	2,213	48**
Missing	5.5	2,802	2,829	27
Heart condition (past 12 months)				
Yes	2.1	8,951	8,132	−819**
No	92.2	2,303	2,337	34**
Missing	5.7	2,962	2,994	32
Migraine (past 12 months)				
Yes	15.0	2,483	2,374	−109**
No	72.3	2,332	2,384	52**
Missing	12.7	3,305	3,266	−39
Hypertension (past 12 months)				
Yes	16.0	4,337	4,168	−169**
No	71.5	1,942	2,003	61**
Missing	12.5	3,163	3,154	−9
Peripheral artery disease (past 12 months)				
Yes	2.5	7,697	7,076	−621**
No	84.9	2,217	2,257	40**
Missing	12.5	3,202	3,182	−20
Asthma or COPD (past 12 months)				
Yes	7.9	4,764	4,586	−178**
No	79.8	2,151	2,188	37**
Missing	12.3	3,140	3,143	3
Psoriasis (past 12 months)				
Yes	2.7	4,017	3,573	−444**
No	84.3	2,305	2,338	33**
Missing	13.0	3,290	3,288	−2
Chronic dermatitis (past 12 months)				
Yes	4.8	2,943	2,798	−145**
No	82.7	2,333	2,366	33**
Missing	12.5	3,264	3,225	−39
Severe/recurrent dizziness (past 12 months)				
Yes	4.1	6,158	5,557	−601**
No	83.3	2,187	2,239	52**
Missing	12.6	3,214	3,191	−23
Severe/recurrent disease of intestines >3 months (past 12 months)				
Yes	4.3	5,731	5,024	−707**
No	83.4	2,204	2,263	59**
Missing	12.3	3,200	3,178	−22
Incontinence (past 12 months)				
Yes	6.3	5,728	5,502	−226**
No	81.1	2,113	2,155	42**
Missing	12.6	3,201	3,170	−31
Arthrosis or arthritis of hip(s)/knee(s) (past 12 months)				
Yes	13.2	4,791	4,581	−210**
No	74.6	1,967	2,027	60**
Missing	12.2	3,110	3,102	−8

(continued)

Table A1. (continued)

Self-reported condition in year 2012	Population frequency, %	Mean spending in 2013, €	Mean predicted spending for 2013 according to the RA model 2016, €	Mean financial result in 2013 under the RA model 2016, €
Chronic inflammation of joints (past 12 months)				
Yes	5.0	5,806	5,488	−318**
No	82.4	2,168	2,207	39**
Missing	12.6	3,193	3,186	−7
Severe/recurrent condition of back (past 12 months)				
Yes	9.9	4,061	3,884	−177**
No	77.7	2,168	2,213	45**
Missing	12.4	3,159	3,141	−18
Severe/recurrent condition of neck/shoulder(s) (past 12 months)				
Yes	9.4	3,767	3,651	−116**
No	78.2	2,209	2,249	40**
Missing	12.4	3,198	3,160	−38
Severe/recurrent condition of elbow/wrist/hand (past 12 months)				
Yes	6.2	4,424	4,304	−120*
No	81.3	2,216	2,249	33**
Missing	12.5	3,218	3,185	−33
Other long-term condition (past 12 months)				
Yes	13.8	5,169	4,799	−370**
No	74.9	1,873	1,961	88**
Missing	11.3	3,202	3,211	9

Note. RA = risk adjustment; COPD = chronic obstructive pulmonary disease.
Statistically significantly different from zero $^{*}p < .05$. $^{**}p < .01$.