Positive and negative effects of risk equalization and risk sharing in regulated competitive health insurance markets

EsCHER Working Paper No. 2022014

November 2022

Richard C. van Kleef, PhD
Mieke Reuser, PhD
Pieter J.A. Stam, PhD
Wynand P.M.M. van de Ven, PhD

Erasmus University Rotterdam
Making Minds Matter

Title

Positive and negative effects of risk equalization and risk sharing in regulated competitive health insurance markets

Authors

Richard C. van Kleef, PhD [a]
Mieke Reuser, PhD [a,b]
Pieter J.A. Stam, PhD [c,d]
Wynand P.M.M. van de Ven, PhD [a]

[a] Erasmus University Rotterdam / Erasmus Centre for Health Economics Rotterdam (EsCHER)
[b] National Institute for Public Health and the Environment (RIVM)
[c] Equalis Strategy & Modeling
[d] School of Business and Economics, VU University Amsterdam

Corresponding author: Richard C. van Kleef (vankleef@eshpm.eur.nl)

Keywords

Health insurance; Risk equalization; Risk sharing; Risk selection; Efficiency

JEL classification

G18; I11; I13; I18

Cite as

Van Kleef, R.C., M. Reuser, P.J.A. Stam & W.P.M.M. van de Ven. (2022). Positive and negative effects of risk equalization and risk sharing in regulated competitive health insurance markets. EsCHER Working Paper Series No. 2022014, Erasmus University Rotterdam. Available from: https://www.eur.nl/en/research/escher/research/working-papers

Abstract

It is well-known that the design of risk equalization and risk sharing in regulated competitive health insurance markets comes with complex tradeoffs between positive and negative effects on efficiency and fairness. This paper provides a comprehensive overview of these effects: in total, we distinguish 22 potential effects, most of which relate to efficiency. The total set of potential effects can be used as an assessment framework for the (re)design of risk equalization and risk sharing schemes. In a second step, we summarize common measures for ex-ante evaluation of risk equalization and risk sharing schemes. The academic literature provides a wide range of measures. In the light of the potential effects, however, some measures are more informative than others, which implies that the choice of measures should be made carefully. Moreover, we find that most measures do not go beyond incentives. The development of more sophisticated measures that incorporate the impact of incentives on behavior is an important step to better predict the effects of alternative risk equalization/sharing designs. In a third step, we discuss how the potential effects are considered in a specific policy context: the Dutch regulated health insurance market. We find that policymakers in the Netherlands do not consider all potential effects. For example, they do not explicitly consider the potential positive effects of risk equalization and risk sharing on the quality of care. To avoid suboptimal policy choices, we recommend policymakers to consider the entire spectrum of potential effects.

Acknowledgement: The authors are grateful to Florian Buchner, Thomas McGuire and the participants of the Risk Adjustment Network meeting in Berlin (September 2022) for helpful comments on earlier versions of this paper. The authors also thank Barry Egberts for his help with summarizing the anecdotal and empirical evidence of risk selection in the Netherlands.

1. Introduction

Many individual health insurance markets are organized by principles of 'regulated competition'. Examples include the mandatory health insurance schemes in Germany, Israel, the Netherlands and Switzerland, voluntary health insurance schemes in Australia and Ireland, and specific sectors in the U.S. such as Medicare Advantage and the state-based Marketplaces that operate under the Affordable Care Act. Economic theory suggests that a well-designed combination of 'competition' and 'regulation' can simultaneously achieve objectives regarding the efficiency and fairness of health insurance systems (Van de Ven et al., 2013).

In terms of regulation, all health insurance schemes mentioned above include premium-rate restrictions. Although these restrictions help achieve fairness objectives, it is well-known that they also generate/exacerbate incentives for risk selection (Van de Ven & Ellis, 2000). To eliminate selection incentives, regulators typically rely on risk equalization and/or risk sharing. By risk equalization we mean a payment system that (re)distributes funds among insurers using indicators of *expected* cost such as age and health. By risk sharing we mean a payment system that (re)distributes funds among insurers on the basis of *actual* cost. Over the past four decades, research has led to major improvements in the design of risk equalization and risk sharing schemes. Algorithms for risk equalization have evolved from simple demographic models to sophisticated morbidity models using health indicators based on (prior) diagnoses, drug prescriptions and other patterns in utilization and/or spending (Ellis et al., 2018). Risk sharing methods have developed from simple modalities of proportional and outlier-risk sharing to more sophisticated forms that target risk sharing payments more directly at losses *net of* risk equalization (McGuire & Van Kleef et al., 2018). Despite these improvements, substantial incentives for risk selection remain (see, for instance, Van Kleef et al., 2019; McGuire et al., 2020; McGuire et al., 2021; Zink & Rose, 2021), calling for further advances.

(Re)design of risk equalization and risk sharing comes with complex tradeoffs between positive and negative effects on efficiency and fairness. Examples of potential negative effects include a reduction of cost control (e.g., when risk equalization and/or risk sharing payments are linked to costs) and a waste of resources due to gaming (e.g., when payments are linked to diagnoses). Examples of potential positive effects include a reduction in risk selection, which enhances various efficiency objectives and helps achieving the intended cross subsidies from healthy enrollees to the chronically ill. Van de Ven et al. (2022) have concluded that "when it comes to the evaluation of a potentially new risk adjuster, a comprehensive analysis needs to be done not only of the negative effects of that new risk adjuster on efficiency (e.g., in terms of gaming), but also of the positive effects on efficiency". The first goal of our paper here is to provide a comprehensive overview of all potential effects of risk equalization and risk sharing (Section 2). Our second goal is to summarize common measures for ex-ante evaluation of risk equalization and risk sharing, and to discuss how these measures relate to the potential effects (Section 3). Our third goal is to discuss how the potential effects are considered in a specific policy context: the Dutch regulated health insurance market for curative care (Section 4). In Section 5 we summarize our findings and discuss the policy implications and directions for future research.

2. Potential effects of risk equalization and risk sharing

To make individual health insurance on a competitive social health insurance market accessible and affordable for high-risk consumers, regulators often implement regulations

such as an open-enrollment requirement and premium-rate restrictions for specified basic health insurance products (see e.g., Van de Ven and Ellis, 2000). By doing so, the regulator intends to create pooling arrangements with implicit cross-subsidies among heterogeneous risks. For simplicity and clarity of our arguments we assume that the regulator requires community rating per product, which means that an insurer must charge the same premium to all insured that enroll in the same health insurance product. In that case the regulator aims to create pooling arrangements such that people with heterogeneous risks pay the same product-related premium. This type of premium regulation is applied in many systems, e.g., the basic health insurance schemes in Germany, the Netherlands, as well as Medicare Advantage in the United States. Other systems include a slightly weaker form of community rating by allowing some limited risk rating according to age and geography, e.g., the basic health insurance in Switzerland and the state-based Marketplaces in the United States. In these cases, the regulator aims to create pooling arrangements such that people with heterogeneous risks pay a product-related premium that may only be risk-adjusted according to the regulation. Although premium-rate restrictions help achieve fairness objectives, they also induce a problem: premiums inherently deviate from the expected costs of insurance contracts, a problem that Newhouse (1996) refers to as 'unpriced risk heterogeneity'. As well-documented in the literature, unpriced risk heterogeneity can lead to risk selection. Inspired by Newhouse (1996), we define risk selection as "actions by consumers and insurers that break or intend to break the pooling arrangements". In Section 1.1, we discuss different types of selection actions as well as their potential effects. Since risk equalization and risk sharing are meant to eliminate selection incentives, reductions of the negative effects of risk selection can be seen as positive effects of risk equalization and risk sharing. These and other effects of risk equalization and risk sharing are summarized in Section 2.2.

2.1 Potential negative effects of risk selection

Our definition of risk selection points at two types of selection actions: 1) actions that intend to break the pooling arrangements (irrespective of whether these actions indeed break the pooling arrangements) and 2) actions that break the pooling arrangements (irrespective of the underlying intention of these actions). Below we discuss both types and their potential negative effects.

2.1.1. Actions that intend to break the pooling arrangements

Without further policy measures (such as risk equalization), community-rated premiums would confront insurers with unpriced risk heterogeneity. On average, young and healthy people would be predictably profitable to insurers while the elderly and chronically ill would be predictably unprofitable. These predictable profits and losses provide insurers with 'incentives to target the young and healthy and deter the elderly and chronically ill' (or, framed differently: 'incentives to break pooling arrangements'). In most health insurance schemes insurers have various instruments for risk selection, either via the design and marketing of insurance products or via other channels such as customer service and supplementary insurance. Examples of selection via product design include structuring coverage in a way that it is relatively unattractive to the elderly and chronically ill, e.g., by not contracting providers who have the best reputation in treating or managing specific diseases. Examples of selection via marketing include selective advertising and providing welcome gifts to healthy applicants. Selection via customer service can be done by not responding (adequately) to queries from people with specific diseases. And selection via supplementary insurance could mean that insurers charge excessive premiums for supplementary insurance

products to groups that are predictably unprofitable for the basic insurance (or that insurers do not accept these groups for supplementary products at all).[1] Below, we describe the potential effects of such actions. It is important to emphasize that these effects are independent of whether the actions are successful or not (i.e., whether they eventually break the pooling arrangement). If these actions indeed break the pooling arrangement, additional potential effects enter the stage, which will be discussed in Section 2.1.2.

A particularly harmful selection action is when insurers structure their coverage (to the extent they are allowed to do so) in a way that their health insurance products are relatively unattractive to high risks (e.g., Cao et al., 2003; Ellis et al., 2013; McGuire et al., 2014). One potential effect of such selection could be that insurance products are not in line with consumer preferences. For example, when insurers choose not to contract with providers who have the best reputation for treating specific diseases, patients might not have access to these providers. Moreover, such actions would threaten the level playing field for providers (in terms of getting contracted by health insurers) and could even discourage physicians and hospitals from acquiring the best reputation in treating or managing specific diseases. That would be an undesirable, inefficient outcome of a competitive healthcare system.

Next to distortions of coverage, negative effects on quality can also result from distortions of customer service. For example, when insurers aim at deterring high-risk enrollees by delaying answers to their letters and emails, letting them wait during phone calls, and otherwise being impolite to them, the quality of customer service will be suboptimal.

Another potential effect of risk selection by insurers is a reduction of cost control. When insurers are confronted with large predictable profits, selection might (be perceived as) a more profitable strategy than improving efficiency in healthcare production. At least in the short run, when an insurer has limited resources available to invest in cost-reducing activities, it may choose to invest in risk selection rather than cost control.

In the presence of unpriced risk heterogeneity, insurers might anticipate adverse selection by consumers and offer a variety of insurance products. For example, these products can differ in terms of coverage, cost sharing and provider network. When one insurer starts offering different insurance products, other insurers must follow to keep attracting predictably profitable people. Such selection-driven product differentiation can seriously reduce the transparency of the insurance market with negative effects on consumer choice and competition.

Finally, investments in risk selection by insurers can be considered a waste of resources because investments purely aimed at attracting low risks through risk selection produce no net benefits to society (risk selection is a zero-sum game among health insurers).

2.1.2. Actions that break the pooling arrangements

[1] The effects of selection via supplementary health insurance can stretch beyond the negative effects on fairness and efficiency of the *basic* health insurance. More specifically, such selection can lead to inefficient pricing of *supplementary* insurance products, and to the unavailability of these products for groups that are predictably unprofitable under the basic health insurance. In the rest of this paper, however, we restrict our consideration of the potential effects of risk selection to effects that directly relate to the basic health insurance.

As mentioned earlier, all potential effects listed in Section 2.1.1 are independent of whether the underlying selection actions are successful or not. To the extent that these actions are successful (i.e., eventually break the pooling arrangements) additional negative effects will occur. Before we start describing these effects, it is important to emphasize that – in addition to 'actions that intend to break the pooling arrangements' – there exists a whole range of 'unintended' actions that can break the pooling arrangements. For example, Beaulieu et al. (2006) have provided empirical evidence of how an HMO that developed a good reputation for treating chronically ill patients attracted disproportionate shares of predictably unprofitable consumers. In general, any correlation between unpriced risk heterogeneity and consumer preferences regarding aspects in which insurance products are allowed to vary (e.g., benefits covered, the level of cost sharing, and the quality of provider networks) can lead to the breaking of pooling arrangements. In other words, the breaking of pooling arrangements can result from many combinations of actions by insurers and actions by insurers, irrespective of the underlying intentions.[2]

Actions that break the pooling arrangements may cause instability in the insurance market, e.g., when low risks permanently switch to lower-priced insurance products (Rothschild and Stiglitz, 1976; Kong et al., 2022). The low-risk individuals may buy (new) lower-priced products that are relatively attractive to them. The premiums for the old products will have to rise as they come to be predominantly bought by high-risk individuals. This may then stimulate some high-risk individuals to switch to these new products too, even if the coverage of these products is suboptimal for them. Consequently, premiums for the new products will increase, which may stimulate the low-risk individuals again to switch to new lower-priced products, even if the restricted coverage of these products is suboptimal for them given their risk aversion. Moreover, since premiums will not only reflect variation in efficiency in healthcare production, but also differences in the insurer's risk composition, these selection-driven premiums distort the consumers' price/quality tradeoff and alter competition on efficiency.

Another inefficiency arising from actions that break the pooling arrangements is the welfare loss due to the potential non-existence of a competitive equilibrium. The continuous exit (bankruptcy) and re-entry of insurers and insurance products come with social costs. Another consequence of actions that break the pooling arrangements is that the cross-subsidies as intended by the regulator are not fully achieved. This may result in unaffordability of health insurance for high-risk individuals. Insurers that specialize in care for undercompensated high-risk patients will have to charge a relatively high premium. In that case high-risk patients can receive good care and good services only if they are able to pay the high premium.

Actions that break the pooling arrangements also distort the level playing field for the insurers. In this context, a level playing field can be defined as a situation in which two insurers who in year *t* have different risk compositions of their insurance portfolio, but are identical in all other aspects (including, for example, their insurance conditions, their provider network, their coverage of out-of-network spending, their cost efficiency, their premium and their financial reserves), have an identical expected financial result in year *t*. A distortion of the level playing field can be considered unfair to the adversely selected insurers. These

[2] Finally, a breaking of pooling arrangements can also be due to historical reasons. For example, in the Netherlands (which, in 2006, introduced one basic health insurance scheme for curative care for the entire population that lives or works in the Netherlands) some insurers find their roots in a public scheme that covered two thirds of the population below a certain income threshold while other insurers find their roots in a private scheme that covered the rest of the population (mostly high-income people).

insurers must charge a higher premium than their competitors, lose market share and may ultimately go bankrupt, even if they are efficient. In addition, it is hard for insurers to set the premium for the next contract period because they do not know the risk profile of consumers who will sort into their insurance products and how many unprofitable high-risk people they must accept during the open enrolment period. This may result in high loading fees to compensate for the risk of being adversely selected, or it may result in the bankruptcy of adversely selected insurers.

Finally, in the absence of an effective insurance mandate, low-risk individuals may not buy health insurance because the community-rated premium far exceeds their expected insurance claims. This will increase the premium for those who do buy health insurance. This may further increase adverse selection, resulting in an upward premium spiral. Consequently, the cross-subsidies as intended by the regulator may not be fully achieved, which may reduce the affordability of health insurance coverage. To the extent that risk averse people remain uninsured, there will be a forgone welfare gain due to suboptimal risk protection. Moreover, uninsured people might not be able to afford expensive treatments in case of serious health problems, especially those with low income (Nyman, 1999).

2.1.3. Potential negative effects of risk selection

In sum, risk selection may lead to negative effects as summarized in Table 1, some of which relate to efficiency while others (also) relate to fairness. Efficiency is a broad concept that encompasses various aspects. Productive (or technical) efficiency of healthcare implies that, given the available resources such as labor and capital, it is not possible to produce more of one good without decreasing the quantity of another good. Productive efficiency relates to static productive efficiency (‘at a certain point in time’) as well as dynamic productive efficiency (‘over time, through investment into production methods and innovation’). Allocative efficiency means that the allocation of production resources optimally reflects consumer preferences. It implies the production of an optimal combination of goods and services that represents the combination that society most desires. Fairness is a broad multi-dimensional concept that covers normative ideas about solidarity, affordability, cross subsidies, equity, justice, and impartiality. Fairness may reflect value judgments that differ among individuals and across countries. We consider fairness towards the consumers, the insurers, and the providers of care.

Table 1. Potential negative effects of risk selection

	Potential effects	Efficiency	Fairness
Potential effects of actions that intend to break the pooling arrangements:			
1	Insurance products are not in line with consumer preferences	x	
2	Absence of a level playing field for providers (since providers with the best reputation in treating specific diseases might not be contracted by insurers)		x
3	Insurers and providers of care do not strive for obtaining the best reputation in treating people with specific diseases who tend to be unprofitable, which may reduce (investments by insurers and providers in) the quality of care.	x	
4	Suboptimal customer service for high-risk enrollees	x	
5	Insurers underinvest in cost control when risk selection is (perceived as) a more effective strategy for making profits	x	
6	Distortion of consumer choice and competition due to reduced transparency of the health insurance market because of product proliferation	x	
7	Waste of resources (i.e., resources used for risk selection do not add any social value since risk selection is a zero-sum game)	x	
Potential effects of actions that break the pooling arrangements:			
8	Consumers choose the 'wrong' insurance product (underinsurance) since premiums of individual insurance contracts are not in line with expected costs	x	
9	Social costs of the absence of a market equilibrium [i.e., excessive exit (bankruptcy) and re-entry of health insurers and health insurance products]	x	
10	Intended income redistribution from low- to high risks is not fully achieved		x
11	Unaffordability of high-value insurance products (which can threaten access to high-quality care and good services, particularly for low-income people)		x
12	Absence of a level playing field for insurers		x
13	Bankruptcy of adversely selected – though potentially efficient – insurers	x	
14	Higher loading fees due to the insurers' risk of being adversely selected	x	
15	Suboptimal risk protection of risk averse (low-risk) people who are not willing to buy health insurance in the absence of an effective insurance mandate	x	
16	Suboptimal risk protection of risk averse (high-risk) people who are willing but not able to buy health insurance in the absence of an effective insurance mandate (due to a lack of cross subsidies from low-risk people)	x	x
17	People who do not buy insurance (see the above-mentioned potential effects 15 and 16) might not be able to afford expensive treatments in case of serious health problems, especially those with low income		x

2.2. Potential effects of risk equalization and risk sharing

In the presence of open enrollment and premium rate restrictions for specified basic health insurance products, risk equalization and risk sharing are expected to reduce most of the

potential negative effects listed in Table 1 (via the reduction of risk selection). The potential positive effects of risk equalization and risk sharing are summarized in Table 2.

Table 2. Potential positive effects of risk equalization and risk sharing (via the reduction of selection)

1	Insurance products are more in line with consumer preferences
2	The playing field for providers is more leveled
3	Better efforts by insurers and providers to acquire the best reputation in treating people with specific diseases who previously were unprofitable, which may improve the quality of care.
4	Better customer service for high-risk enrollees
5	More investments by insurers in cost containment since the potential returns on risk selection decrease
6	A reduction of selection-driven product proliferation which increases transparency on the health insurance market and thereby enhances a value-for-money consumer choice and competition on efficiency.
7	Less resources spent on selection activities
8	Consumers increasingly choose the 'right' insurance product as premium differences between high- and low-value products are less distorted by differences in risk composition across insurance products
9	More stability in the insurance market (resulting in a reduction of the social costs of the absence of a market equilibrium in terms of excessive exit and re-entry of insurers and insurance products)
10	Intended income redistribution from low- to high-risk people more fully achieved
11	Increased affordability of high-value insurance products (which improves access to high-quality care)
12	The playing field for insurers is more leveled
13	A lower chance of bankruptcy of (efficient) insurers due to adverse selection
14	Lower loading fees due to a reduction of the risk of adverse selection

If there is no effective mandate to buy health insurance, the impact of risk equalization and risk sharing on the last three negative selection effects in Table 1 (number 15-17) depends on how risk equalization and risk sharing payments are financed (Van Kleef et al., 2018a). If these payments are financed 'internally', i.e., via the premiums (as e.g., in the ACA marketplaces in the United States and in the voluntary health insurance markets in Ireland and Australia), the premium for low-risk people might increase. More specifically, when low-risk people are concentrated in specific insurance products (e.g., with a high-deductible), risk equalization and risk sharing are likely to drive up the premiums of these products. This may stimulate a new group of low-risk people not to buy health insurance and exacerbates the negative selection effects 15-17 in Table 1. 'External' financing of risk equalization and risk sharing payments, e.g., via taxes or mandatory contributions, will reduce the premium, both for high-risk and low-risk people. The reason is that external risk equalization and risk sharing function as a subsidy to the market, resulting in a decrease of premiums (assuming insurers pass through this subsidy to consumers by lowering their premiums). This may stimulate some uninsured low-risk people to buy health insurance, which reduces the negative selection effects 15-17 in Table 1. In the case of a mixture of internal and external funding of

the risk equalization and risk sharing, the effect depends on the weights given to the internal and external funding.

In practice, the positive effects of risk equalization and risk sharing come with a price. Ideally the risk adjusters used in the risk equalization should be valid, reliable, and non-manipulable, and the required data must be available at socially acceptable costs. Because of these conditions, it is likely that the goal of risk equalization can only be achieved at a price, such as high costs to collect the required data or including risk factors that decrease the incentives for cost containment. State-of-the-art risk equalization schemes have diagnoses/cost-based risk adjusters which can reduce the insurers' incentives for cost containment, because a reduction of healthcare utilization or healthcare expenses may result in lower future equalization payments. Risk equalization schemes may also lead to perverse incentives. For example, insurers and/or the contracted providers may provide unnecessary services to code a diagnosis, upcode diagnoses to more serious conditions, (fraudulently) change the diagnostic coding, or (fraudulently) distort information reported to the regulator that is used for payment purposes.

Equalization payments based on prior diagnoses might also reduce incentives for prevention, (e.g., they may discourage lifestyle interventions for – potential – diabetes patients), because if an insurer improves the health status of its enrollees by good quality care and effective prevention, this may result in lower future equalization payments.[3]

Like risk equalization, risk sharing also comes with a price (Newhouse, 1996). The ex-post cost-based payments to the insurers reduce the insurers' incentives for cost containment (i.e., quality improvement or/and cost-reduction efforts) and prevention. The extent of this negative effect depends on the amount and the funding of the ex-post compensations and on other aspects of the specific form of risk sharing (e.g., Van Barneveld et al., 2001; Van Kleef & Van Vliet, 2022).

The negative effects of risk equalization and risk sharing are summarized in Table 3.

[3] However, as Van de Ven and Ellis (2000) indicate, a counter argument is that improved health not only reduces future revenues, but also future expected costs. Furthermore, if an insurer effectively reduces the incidence of e.g., heart diseases, it fully benefits from not having the high first-year expenses related to these diseases. (This argument does not hold in case of concurrent risk equalization.) In addition, the insurer fully benefits from not having expenses related to preventable transitory health problems for which the equalization payments are not adjusted (e.g., fever and flue). Nevertheless, it is true that an insurer bears the full costs of health-improving activities and preventive services such as smoking cessation, weight loss, and nutritional guidance, while it may lack a part of the future returns. Whether in practice these incentives override the consumer preferences and the professional ethics of the providers, remains an empirical question (see also Kanters et al., 2013). It is interesting to note that in free competitive insurance markets insurers often use such risk factors that potentially have the above-mentioned negative effects, while nevertheless such competitive markets are generally considered to be efficient. Van de Ven et al. (2022) give a possible explanation why such risk factors are a (potential) problem in a community-rated market with risk equalization and not in a free market with risk rating.

Table 3. Negative potential effects of risk equalization and/or risk sharing

18	Cost of data collection and preparation
19	Reduction of cost control
20	Unnecessary services to code a diagnosis
21	(Fraudulent) change of the diagnostic code (e.g., upcoding of diagnoses to more serious conditions) and other distortion of the information reported to the regulator
22	Reduction of prevention

In sum, risk equalization and risk sharing have both positive and negative potential effects that require a complex trade-off. Tables 2 and 3 can be used by policymakers as a qualitative assessment framework for making decisions on (re)design of their risk equalization and/or risk sharing scheme. It is important to note that the occurrence and size of the effects in Tables 2 and 3 depend on the institutional context and the precise form of risk equalization and risk sharing, e.g., internal or external funding, and prospective or concurrent risk equalization (see e.g., Van de Ven and Ellis, 2000). Moreover, it is important to note that health insurance systems differ substantially in terms of the 'scope for selection', resulting in large differences in potential actions and effects of selection. For example, the scope for selection is broader as 1) individual insurers have more flexibility in the use of tools to manage costs and/or quality of care, 2) each insurer may offer multiple health insurance products (that fulfill the regulation), resulting in more consumer choice options, and 3) legal barriers for insurers to enter the market are absent.

3. Measures for ex-ante evaluation of risk equalization and risk sharing

This section summarizes the measures for ex-ante evaluation of risk equalization and risk sharing that have been developed and used in the literature and discuss how these measures relate to the potential effects discussed in Section 2.2. We first summarize measures related to risk selection (Section 3.1) and then discuss measures related to gaming and cost control (Section 3.2). We focus on what we call 'ex-ante' measures, i.e., measures used to *predict* incentives, actions and/or effects under risk equalization and risk sharing schemes. Papers in this stream of literature include studies on the development and evaluation of risk adjuster variables (e.g., Ash et al., 2000; Pope et al., 2004; Ellis et al., 2022), studies on the evaluation of 'payment fit' under existing payment systems (e.g., Stam et al., 2010a; Montz et al., 2016; Withagen-Koster et al., 2018; Van Kleef et al., 2019) and studies on the design and evaluation of risk sharing modalities (e.g., Van Barneveld et al., 2001; Beck et al., 2020; McGuire 2020 and 2021; Van Kleef & Van Vliet 2022). Given their ex-ante nature, these studies differ from the stream of literature on ex-post measurement of *actual* actions and effects, e.g., in terms of risk selection and gaming (Bauhoff, 2012; Van de Ven et al., 2017; Han & Lavetti., 2017; Lavetti & Simon., 2018; Geruso et al., 2019; Geruso & Layton, 2020; Shepard, 2022).

3.1. Ex-ante measures related to risk selection

Risk equalization and risk sharing are meant to eliminate selection problems by compensating insurers for 'unpriced risk heterogeneity', i.e., variation in *expected* costs of insurance

contracts that is *not* allowed to be explicitly reflected in premium variation. When risk equalization and/or risk sharing perfectly compensate for unpriced risk heterogeneity, risk selection will be absent (since pooling arrangements will stay intact even when healthy and unhealthy consumers sort into different insurance products). Below we discuss some common measures that have been used to quantify the extent to which risk equalization and risk sharing compensate for unpriced risk heterogeneity. In line with Section 2, we make a distinction between ex-ante measures that relate to 'actions that intend to break the pooling arrangements' (Section 3.1.1.) and ex-ante measures that relate to 'the breaking of pooling arrangements' (Section 3.1.2).

3.1.1. Ex-ante measures related to actions that intend to break the pooling arrangements

In the academic literature, many ex-ante measures have been developed and applied that relate to selection by insurers. For a comprehensive overview, see Van Veen et al., (2015) and Layton et al. (2017). We categorize existing measures into three groups: 1) measures of statistical fit, 2) measures of incentives and 3) measures of expected actions and their effects. Below, we describe these categories and highlight common measures in each category.

3.1.1.1. Measures of statistical fit

The most applied measure in ex-ante evaluations of risk equalization models is the R-squared. This measure – which comes from the field of statistics – indicates the proportion of variance in medical spending that can be explained by the independent variables (i.e., risk adjusters) of the risk equalization estimation model. Recent studies have upgraded the R-squared to 'Payment System Fit' (Geruso & McGuire, 2016), a measure that shows the proportion of variance in medical spending compensated for by the entire payment system (which – in addition to risk equalization – can consist of risk sharing and/or premiums).[4] Although the R-squared and PSF are informative, the link between these measures and the effects listed in Section 2 of this paper is ambiguous. The reason is two-fold. First, the R-squared and PSF include a quadratic weighting of errors. Although large errors may be more problematic than small errors, it is not obvious that quadratic weighting is better for evaluations of risk equalization than alternative forms of weighting. To overcome this shortcoming, some studies have (also) applied linear measures of statistical fit such as Cumming's Prediction Measure (CPM) and the Mean Absolute Prediction Error (MAPE). Second, the R-squared and PSF (as well as the CPM and MAPE) summarize gaps (errors) between individual-level predicted costs (or: revenues) and *actual* costs. Selection incentives, however, result from gaps between revenues and *expected* costs. This is an important shortcoming since the lion's share of individual-level variance in spending might not be *predictable* and thus cannot be anticipated on by insurers. The relevance of this shortcoming can be illustrated with the following example: a payment system with 50% proportional risk sharing has a PSF of 0.75 and a CPM of 0.50 while systems with state-of-the-art prospective risk equalization have a PSF and CPM that are much lower, somewhere in the range of 0.30-0.35. Despite the lower PSF and CPM,

[4] PSF has been used in recent papers such as Schmid et al. (2016) in an evaluation of the health insurance payment system in Switzerland and McGuire et al. (2020; 2021) in comparative studies on the health insurance payments systems included in Germany, The Netherlands, and the State-based Marketplaces in the U.S.

however, these risk equalization models are likely to better mitigate risk selection incentives than 50% proportional risk sharing.[5]

3.1.1.2. Measures of incentives

Many studies have recognized the shortcomings of the R-squared, CPM, MAPE and PSF and (also) apply measures that more directly relate to *incentives* for insurers to engage in risk selection. Common measures of incentives include under/overcompensation and predictive ratios for groups of interest (e.g., Stam et al., 2010b; Montz et al., 2016; Withagen-Koster et al., 2018; Van Kleef et al., 2019). Whereas under/overcompensations show the monetary value of the difference between predicted costs and actual costs for subgroups, predictive ratios show the ratio of predicted costs and actual costs for subgroups. If the groups are large enough, the actual average costs per person in such a group can be interpreted as the 'expected costs' for a person in that group. If these measures are calculated for groups that are (potential) targets of specific selection actions, they meaningfully indicate the incentives for insurers to engage in these actions. Most studies that calculate group-level under/overcompensations or predictive ratios for year *t*, identify groups in data from year *t*-1. For example, Van Kleef et al., (2019) calculate under/overcompensation for groups identified in a health survey from *t*-1 and Van Kleef et al. (2020) identify groups using diagnoses from *t*-1. By identifying groups with information from a prior period, these measures inherently isolate predictable (or systematic) spending variation from random spending variation (assuming these groups have a sufficient size). Alternative measures to isolate predictable spending variation from random spending variation have been applied by Lamers (2001) and Stam et al., (2010a) who rely on individual-level measures – like the statistical measures discussed above – but with a different benchmark: instead of comparing revenues to observed costs these studies compare revenues to *expected* costs. Gaps between predicted and expected costs indicate 'unpriced risk heterogeneity' and thus selection incentives.

Other meaningful measures of incentives are 'predictiveness and predictability'. Predictability refers to the extent to which use of a specific service (e.g., home care) in year *t* is predictable. Predictiveness refers to the extent to which spending on that service in year *t* correlates with the overall (un)profitability in year *t*. Together, predictiveness and predictability indicate the insurers' incentives for service-level distortion, i.e., the incentives to deviate from the socially optimal allocation of resources across medical services (Ellis & McGuire, 2007). More specifically, when use of services (e.g., home care) is to some extent predictable and spending on that service negatively correlates with overall profitability, insurers face incentives to underprovide that service. Where group-level under/overcompensation and predictive ratios indicate incentives for insurers to select in favor or against specific groups, predictiveness and predictability more directly indicate incentives for under/overproviding specific medical services. Alternatively, such incentives can be indicated by under/overcompensation for users of specific services in the prior year, such as the group that used home care in the prior year.

3.1.1.3. Measures of expected actions and their effects

[5] Van Kleef et al. (2019) find that the prospective risk equalization model used in the Netherlands in 2016 has a PSF and CPM of 0.3 and reduces predictable profits/losses for subgroups of interest by >80%. Although the authors do not simulate the effects of a payment system that solely consists of 50% proportional risk sharing, it can be analytically argued that such a system comes with a PSF of 0.75, a CPM of 0.50 and a reduction of predictable profits/losses for subgroups of interest by 50% (assuming that premiums are community-rated).

Although group-level under/overcompensation and predictiveness and predictability relate more directly to incentives for selection by insurers than measures of statistical fit such as the R-squared and CPM, they do not capture the 'expected effects.' To go beyond incentives, one must have a clear idea about how incentives translate into 'actions' (whether to engage in a specific selection action or not). Only a few studies have explored going beyond incentives. The most prominent stream of literature on this topic are the papers on 'service-level distortion' (Frank et al., 2000; Glazer & McGuire, 2002; Layton et al, 2018). These papers rely on a model of insurer behavior. By making assumptions on the objective of insurers (profit maximization), the level on which insurers 'take action' (decisions on how to allocate premium revenues to healthcare services) and the insurers' expectation of consumer behavior (how consumers value service-level allocations), this model allows predicting how much insurers will spend on specific services under a given payment system. Since the payments that insurers receive directly feed into the profitability of enrollees, alternative payment systems will lead to different service-level allocations. When the 'socially optimal' service-level allocations are known, this model allows for indicating the 'welfare loss' under payment model *m* as the discrepancy between the predicted service-level allocation under *m* (in equilibrium) and the optimal allocation.

Although the model of profit maximization provides the most comprehensive ex-ante measure that allows for predicting effects of selection by insurers under risk equalization and risk sharing, it has not been commonly used in policy research so far. A meaningful application requires valid estimations (or at least plausible assumptions) on insurer behavior in a specific setting. In many settings, such estimations might not be available. Moreover, the original model of profit maximization has been designed for one type of selection action (service-level distortion). In a context with many other potential selection actions, this model needs modification (or additional models must be developed for the other potential actions).

Some studies follow a more pragmatic approach for going beyond incentives by applying some form of non-linear and/or asymmetric weighting of under/overcompensations. For example, Van Barneveld et al. (2000) ignore small under/overcompensations based on the assumption that insurers are unlikely to act on small under/overcompensations given the costs of risk selection. Withagen-Koster et al. (2020) give more weight to undercompensation of unhealthy groups than undercompensation of healthy groups based on the assumption that undercompensation of unhealthy groups is more likely to result in quality skimping and therefore comes with larger (or worse) welfare losses than undercompensation of healthy groups. A meaningful application of such non-linear and asymmetric weighting, however, requires consensus about 1) the groups or services of interest and 2) the weighing of under/overcompensations for these groups/services given the likelihood of selection actions and the potential effects of these actions.

3.1.2. Ex-ante measures related to the breaking of pooling arrangements

If all insurers for all their products are equally successful with their selection actions, the distribution of low- and high-risk individuals might be roughly similar across products. Under such circumstances, the potential effects of risk selection are 'limited' to the effects 1-7 in Table 1. However, as soon as low-risk and high-risk people start sorting into different products, additional effects will enter the stage. These effects result from differences in mean per person net spending (i.e., spending net of risk equalization and/or risk sharing) across insurance products. For products with disproportionate shares of profitable people, mean per person net spending will be relatively low. For products with disproportionate shares of

unprofitable people mean per person net spending will be relatively high. Such sorting can lead to the negative effects 8-14 in Table 1. Risk equalization and risk sharing can mitigate or eliminate these problems by compensating insurers for differences in risk composition across products.

One measure to evaluate the extent to which a payment system compensates for (potential) risk variation across products has already been discussed in Section 3.1.1: group-level under/overcompensation. This measure does not just indicate the incentives for insurers to select in favor or against specific groups (Section 3.1.1.2) but also the under/overcompensation of insurance products under *extreme* sorting patterns. For example, Van Kleef et al. (2019) find that the Dutch risk equalization model of 2016 overcompensates the group of people who reported a (very) good health in the prior year by 187 euro per person per year and undercompensates the complementary group by 512 per person per year. These figures imply that under the strong assumption that the two groups perfectly sorted into different products, the mean costs of these products (net of risk equalization) would differ by 699 euro per person per year. Such perfect sorting, however, is unlikely and should be regarded as one end of the spectrum, with the other end being 'no sorting'. Researchers can easily simulate product-level outcomes for sorting patterns in between these two extremes by simulating outcomes for 'hypothetical' portfolios with disproportionate shares of healthy/unhealthy enrollees.

Measures of group- or hypothetical-portfolio-level under/overcompensation indicate the extent to which pooling arrangements potentially break under a risk equalization/sharing scheme, but do not directly predict the expected effects. For a meaningful prediction of effects, a model is needed on insurer and consumer behavior. On the insurer side, such a model needs valid estimations (or at least plausible assumptions) on how risk equalization and/or risk sharing payments translate into premiums. More specifically, researchers must have a clear idea of how under/overcompensation of a population with a certain insurance product will be reflected in a higher/lower premium for that product, ceteris paribus. Only with perfect competition, product-level under/overcompensations will be fully reflected in premiums. Due to market power and choice frictions, however, competition in health insurance markets tends to be imperfect. As a result, the 'pass through' of under/overcompensation into premiums may incomplete. Indeed, Cabral et al. (2018) find a pass-through rate of about 50% in Medicare Advantage. Since the level of competition is likely to differ across countries and markets, estimations of insurer behavior in one setting cannot be extrapolated to other settings. Such estimations need to be derived for each setting separately. On the consumer side, prediction models of consumer sorting need valid estimations of demand for (specific) insurance (products) and the correlation of demand with expected costs. With such a model, researchers can predict the sorting equilibrium in health insurance markets and the welfare loss resulting from inefficient sorting due to adverse selection. For the workings and application of such models see Einav et al. (2010), Einav and Finkelstein (2011), Bundorf et a. (2012) and Hackman et al. (2015). With a well-designed model of insurer and consumer behavior, researchers can predict and evaluate the sorting patterns under alternative risk equalization and/or risk sharing designs. See, for instance, Handel et al. (2015), Saltzman (2021) and Geruso et al., (2021). In terms of welfare effects, all these studies focus on the 'forgone welfare gain' from underinsurance, either in the case of not buying insurance or in the case of buying too little insurance. These studies do not explicitly consider the other potential effects of the breaking of pooling arrangements.

3.2. Incentive measures for cost control and gaming

Risk sharing systems provide insurers with payments on the basis of *actual* cost. Such a link between costs and payments reduces the incentives for insurers to control costs, or in the terminology of Geruso & McGuire (2016) such a link reduces the 'power' of the payment system. For example: with 10% proportional risk sharing, an increase (decrease) in the insurer's cost by 1 euro implies an increase (decrease) of the insurer's risk sharing payment by 10 cents, implying a power of 0.9 [power can range from 0 (no power) to 1.0 (maximum power)]. With other forms of risk sharing, such as risk corridors (sharing of product-level profits and losses outside a bandwidth) or outlier-risk sharing (sharing of individual-level costs above a threshold), power is somewhat harder to quantify. A pragmatic approach for indicating power under these risk sharing systems is to just look at the fraction of overall revenues that is allocated via risk sharing as done by McGuire et al. (2020; 2021). A more sophisticated approach has been applied by Van Kleef & Van Vliet (2022) and Withagen-Koster (2022) who simulate the effect of a price increase for an average risk portfolio on the payments from a high-risk pool.

Like risk sharing, risk equalization can create a link between costs/utilization and payments too. If risk equalization is solely based on exogenous risk adjusters (i.e., risk adjusters that cannot be influenced by insurers, such as age and gender) it has a power of 1.0. But if risk equalization is (partly) based on endogenous risk adjusters (i.e., risk adjusters that can be influenced by insurers, such as indicators based on prior spending and diagnoses) power falls below 1.0. In the spirit of Geruso & McGuire (2016), the power of a risk equalization system can be simulated by virtually removing (or adding) diagnoses and calculating the change in payments relative to the change in costs (i.e., the costs of the treatments from which these diagnoses are derived).

When it comes to gaming, incentives for insurers to upcode or explicitly induce demand of unnecessary care can be indicated by comparing the 'costs' of a treatment that leads to a risk adjuster flag and the payment weight associated with that flag. This approach is applied by Van Kleef & Van Vliet (2011) who compare the cost of various medical devices for patients (such as prosthesis and tube feeding equipment) with the insurers' costs of these devices. For all fourteen devices included in the Dutch risk equalization model of 2008, the payment weight far exceeds the insurers' costs, implying substantial incentives for gaming.

The aforementioned measures of cost control and gaming are pure incentive measures and do not predict the expected effects. For the prediction of effects, researchers must know how incentives translate into behavior which requires a clear idea of the 'actions' that insurers can and will take in terms of cost control and gaming. For example: if insurers have no instruments to control cost, there is no need to worry about the power of a payment system. In addition, researchers must have valid estimations or plausible assumptions about the link between incentives and behavior. For example, it is questionable whether the link between power and cost control efforts is linear. It could well be that above a certain level of power insurers might actively invest in cost control but that below that level they will abruptly stop these investments, implying a non-linear link between power and cost control efforts. As far as we know, no study has gone beyond indicating incentives for cost control and gaming by predicting effects.

3.3. Conclusion

This section has provided an overview of ex-ante measures used in the academic literature to evaluate risk equalization and risk sharing schemes (or to evaluate selection incentives more generally). Based on this overview, two important observations can be made. First, in the light of the potential effects of risk equalization and risk sharing, some measures are more informative than others. This implies that the choice of measures should be made carefully, i.e., should link to the potential effects that are relevant in the context of interest. Measures of statistical fit, such as the R-squared, CPM, MAPE and PSF, are hardly informative since the link between these measures and the potential effects of risk equalization and risk sharing is rather weak. Measures of under/overcompensation for groups of interest are much more informative, either for indicating the incentives for insurers 'to break pooling arrangements' or for indicating the extent to which 'the pooling arrangement can break' under a risk equalization/sharing scheme. Second, measures of under/overcompensation have their limitations too since they do not directly indicate the expected effects. In the literature only two types of measures have been developed that allow for prediction of effects by using economic models of how incentives (are expected to) feed into behavior: 1) the measures of service-level distortion that originate from the early work by Frank et al. (2000) and 2) the measures of forgone welfare gain due to underinsurance of low-risk individuals (e.g., in the spirit of Einav et al. (2010)).

4. Which potential effects are considered in policy and how? The case of the Netherlands

This section summarizes which potential effects of risk equalization and risk sharing are considered in a specific policy context: the Dutch basic health insurance for curative care. This health insurance scheme is based on the model of regulated competition. Important aspects of competition in this scheme include a free consumer choice of health insurer (which generates competition among insurers) and a toolkit for insurers to improve efficiency of care (such as the possibility to selectively contract with providers of care, which generates competition among providers). Important regulatory features include an insurance mandate for consumers, an open enrollment requirement for insurers, community-rating per insurance product and a system of risk equalization and risk sharing. For two reasons, the Dutch scheme provides an interesting context in the light of this paper. First – in theory – the entire spectrum of effects discussed in Section 2 is potentially relevant in the Dutch scheme, except effects 15-17 due to the presence of an effective insurance mandate. Second, over the past three decades the Dutch have developed an explicit framework for ex-ante evaluation of (modifications of) the risk equalization model, which allows us to examine which potential effects are explicitly considered and how.

Below, we first describe the risk equalization policy and research cycle in the Netherlands (Section 4.1). In a next step, we discuss the effects that are considered quantitatively (Section 4.2) and qualitatively (Section 4.3). Finally, we briefly discuss how other signals, for example, from the media or ex post evaluations can influence policymaking (Section 4.4).

4.1. Policy & research cycle in the Netherlands

The annual policy and research cycle for maintaining and improving the risk equalization model starts with a research agenda and ends with a recommendation to the Minister of Health on the preferred modifications. (In this analysis we do not consider the final decision by the Minister or the legal procedure to enact the modification.) The policy and research cycle in the Netherlands is rather unique, in the sense that there is a strongly committed Expert Committee supervising all stages of the research cycle. This Expert Committee consists of about 70 representatives from research institutes, health insurers as well as governmental bodies like the Ministry of Health, the Ministry of Finance, and the Healthcare Institute (which is responsible for the implementation of the risk equalization model and the calculation of transfers and payments to insurers). The chairperson and the secretary come from the Ministry of Health.

The annual cycle for the risk equalization model of year *t* starts in July of year *t*-2 with the formulation of a research agenda that is sent to Parliament in summertime. The research projects start in the fall and are carried out in two phases. In the first phase, separate research projects are conducted to explore potential model modifications. This often entails refinement and updates of existing risk adjusters (e.g., an update of the set of diagnoses used for the risk adjuster 'Diagnostic Cost Groups') but can also be about new risk adjusters (e.g., a risk adjuster for yes/no 'giving birth' in the payment year, as developed in 2022 for implementation in the model of 2023). These separate research projects are supervised by a selection of the Expert Committee. Once these projects are completed, the second phase starts, i.e., a simulation of the effects of all proposed model modifications from the separate projects combined. This second phase takes place in the summer of year *t*-1 and uses the most recently available data. Based on the results of this research, the Expert Committee advices the Minister of Health on the preferred (combination of) modifications of the risk equalization model for year *t*.

To evaluate the effects of (potential) model modifications the Expert Committee developed an Evaluation Framework (WOR 871). This framework functions as a guideline, both for researchers conducting risk equalization research in the policy-cycle, as well as for the Expert Committee to objectivize model evaluations and recommendations for modification. This Evaluation Framework consists of quantitative measures for ex-ante evaluation of alternative risk equalization models as well as more qualitative measures or arguments to consider in decision making, like cost control incentives, complexity, validity, and data-reliability.

4.2. Risk equalization and risk sharing effects considered quantitatively

The Evaluation Framework includes a series of specific quantitative measures to assess the statistical fit between predicted spending (generated by a risk equalization model) and actual spending (in the dataset on which that model is estimated).[6] The typical procedure is to compare the statistical fit of the current risk equalization model, as a benchmark, to the statistical fit of one or more alternative models with altered risk adjusters.

4.2.1. Quantitative measures to evaluate predictive performance

Since the introduction of risk equalization in the Netherlands in 1993, the R-squared has been one of the most-used measures to evaluate the performance of the risk equalization model. All

[6] Research on the calibration, evaluation and modification of the Dutch risk equalization model applied in year t is typically based on a dataset with spending from year t-3 and individual risk characteristics from the period t-8 to t-3 (for a detailed description see Van Kleef et al., 2018b).

research projects have reported the R-squared on the level of individuals and the level of insurer portfolios. Since 2015 the Cumming's Prediction Measure (CPM) has also been reported in risk equalization research reports. Another common measure is the mean absolute prediction error (MAPE), which is reported on the individual level, insurer level and subgroup level (across all combinations of risk adjusters included in the risk equalization model, in 2017: 1.85 million groups in total). In addition, the range of the mean per person prediction error per health insurer is reported. For each individual insurer the mean per person prediction error is anonymously reported, which can be interpreted as the under/overcompensation of that insurer.

In 2017, a risk equalization symposium led to a revision of the Evaluation Framework. With input from researchers, health insurers and other experts, the Evaluation Framework was extended and elaborated. One of the important differences was the inclusion of mean prediction errors of subgroups. Position papers of the symposium, several academic articles as well as debates in the parliament expressed concerns about over- and undercompensation of specific groups of consumers, especially undercompensation of people with a chronic disease (VWS, 2016). Literature suggests significant and structural under-, and overcompensation of groups defined by characteristics that are not explicitly included as risk adjusters in the model, for example groups identified in a health survey (Van Kleef et al., 2019) or groups of individuals with chronic diseases (Van Kleef et al., 2020). Specific subgroups that were added to the Evaluation Framework include 1) consumers in the bottom-15% of the spending distribution three years ago and 2) those in the top-15% in the spending distribution three years ago.

Table 4: Quantitative measures included in the Dutch Evaluation Framework for evaluating the risk equalization model (source: WOR 871)

Level of measurement	**Quantitative measure**
Individuals	R-squared
	Cummings Prediction Measure
	Mean Absolute Prediction Error (MAPE)
Subgroups	Mean weighted absolute prediction error over all combinations of risk adjusters included in the risk equalization model (weighted with the number of insured years in each combination)
	Mean per person prediction error for somatic care for the following groups: • Bottom-15% of spending distribution for somatic care three years ago • Top-15% of spending distribution for somatic care three years ago • Spending for home care in the prior year > 0
	Mean per person prediction error for mental care for the following groups: • Bottom-25% of spending distribution for mental care in the prior year • Top-1% of spending distribution for mental care in the prior year • Top-3% of spending distribution for mental care in the prior year • Top-5% of spending distribution for mental care in the prior year • Spending for mental care in the prior year > 0 • Spending for mental care in the prior year = 0
Insurer portfolios	R-squared
	Mean weighted absolute mean per person prediction error over insurer portfolios (weighted with the number of insured years per portfolio)
	Range (difference between min and max) of mean per person prediction error
	Mean per person prediction error (anonymously)
	Mean weighted absolute change in prediction error when moving from model A to B (weighted with the number of insured years per portfolio)

Another measure that is often calculated in ex-ante evaluation of the risk equalization model is the sum of funds that the model redistributes across individuals in the total population. This amount (and development is this amount over time), was sometimes used in letters to the Parliament as a measure of performance of the risk equalization model.

It is interesting to note that six measures in the Dutch Evaluation Framework are not considered in academic research (see Section 3). This holds for the first measure at the group level and all measures at the insurer level in Table 4. All these six measures are hard to interpret and are invalid measures for quantifying selection incentives. A problem with the mean weighted absolute prediction error over all combinations of risk adjusters is that it can substantially underestimate incentives for risk selection. For example, adding interaction terms to a poorly performing risk equalization model (e.g., a model with just two dummy variables, one for young/old and one for male/female), the mean weighted absolute prediction error over all combinations of risk adjusters will be zero, while substantial selection incentives remain.
A problem with the measures at the insurer level is that the outcomes of these measures heavily depend on the distribution of risk types across insurers' portfolios and on the cost structure in these portfolios (e.g., in terms of volume and type of healthcare utilization and prices of care). For example, consider the worst risk equalization formula with for each person the predicted expenses equal to the mean per person expenses. Such a risk equalization model has maximum incentives for risk selection. Nevertheless, if the risk composition and

cost structure are identical across the insurers' portfolios, the R-squared is 1.0 and the MAPE is zero. So, the R-squared and the MAPE at the insurer level are no valid indicators of selection incentives. Also, under the same assumptions, the range of the mean per person prediction error (= financial result) in the dataset used for evaluating the risk equalization model (with spending from year *t*-3, see footnote 6) is not a good indicator of a level playing field in year *t*. In addition, in the Netherlands about 6 percent of the population switch during the annual open enrollment period. So, in year *t* (i.e., 3 years later) the risk composition of insurers' portfolios might have substantially changed. And in the case of 'perfect' risk equalization, the (range of the) mean per person prediction error reflects the different cost structures of the insurers and may not be zero.

4.2.2. Quantitative measures for cost control

In addition to the measures mentioned in Table 4 (which evaluate the fit between predicted and actual spending) negative effects of risk equalization and risk sharing on cost control have always played an important role in the evaluation of the risk equalization model and in decisions for model modifications. The Evaluation Framework also pays a lot of attention to incentives for cost containment, both in choosing the risk adjusters as well as in the design of risk adjusters.

The effects of risk equalization and risk sharing on cost control are mainly evaluated by qualitative considerations. The framework states, for instance, that a risk adjuster based on health status information is always preferred over a risk adjuster based on actual (prior) spending. The only quantitative measure included in the Evaluation Framework is the so-called 'repayment ratio', which is closely related to the power measure mentioned in section 3. This measure takes the ratio of the 'payment weight for risk adjuster *k*' over the 'mean per person cost of the treatment (or pharmaceutical) that leads to a score on risk adjuster *k*'.

4.3. Risk equalization and risk sharing effects considered qualitatively

Some of the criteria in the Evaluation Framework are not evaluated via quantitative measures but via qualitative considerations. Apart from the negative effects of risk equalization and risk sharing on cost control incentives, the considerations mainly relate to implementation issues like complexity and transparency of the model and the validity and measurability of the risk adjusters. As mentioned above, the effect of risk equalization and risk sharing on cost control are mainly considered in a qualitative way. These considerations have led policymakers to prefer health-based over cost-based risk adjusters, and prospective over concurrent morbidity indicators. One clear and substantive example of how gaming (or upcoding) incentives play a role in risk equalization design are the thresholds of (e.g., 180) defined daily doses applied to the Pharmaceuticals Cost Groups to prevent 'gaming', e.g., by prescribing small doses of drugs.

These considerations of cost control concerns have been fundamental in the discussion on ex-post risk sharing. The political believe has been that financial responsibility for insurers promotes cost control, which contributes to sustainable health spending. Therefore, the coalition agreement of 2012 aimed to reduce risk sharing (VVD & PVDA, 2012). Explicit quantitative measures to assess incentives or effects of risk sharing were never analyzed or used.

4.4 Other considerations on risk equalization and risk sharing effects

So far, this section has focused on the potential effects considered in the formal policy/research cycle for the design of risk equalization and risk sharing in the Netherlands. Policy debates and decisions, however, are also affected – either directly or indirectly – by anecdotal and empirical evidence of insurer and consumer behavior in the health insurance market. One example of such evidence is an ex-post evaluation of the health insurance market by the Dutch health Authority and the Authority for Consumers & Markets (NZA/ACM, 2018). These authorities evaluated the distortion of consumer choice and competition in the Dutch health insurance market. The abundance of nearly-similar insurance products offered on the market and the lack of transparency of product features distort consumer choice, which potentially results in consumers choosing suboptimal products or consumers not switching at all (ACM/NZA 2018). This effect was described in Table 1 as 'Distortion of consumer choice and competition due to reduced transparency of the health insurance market because of product proliferation". The results of this report were picked up by the media and the Parliament. The transparency of the health insurance market has become an important policy goal (VVD, CDA, D66 & CU, 2017). Table 5 describes some other signals of risk selection that have influenced the policy debate.

The signals of risk selection in Table 5 highlight the relevance of considering the entire spectrum of potential negative effects of risk selection listed in Table 1 (except 15-17 due to the insurance mandate in the Netherlands). More specifically, signals 1-4 in Table 5 indicate that insurers seem to be responsive to selection incentives, which is in line with a growing international literature on selection by insurers (e.g., Bauhoff, 2012; Geruso et al, 2019; Lavetti & Simon, 2018; Han & Lavetti, 2017). Signals 5-8 indicate that – to some extent – low-risk and high-risk consumers sort into different insurance products, which is in line with international research too (e.g., Trottmann et al., 2012; Newhouse et al., 2015; Sheppard, 2022). In other words, the Dutch regulated health insurance market seems to be subject to 'actions that intend to break the pooling arrangement' as well as 'actions that break the pooling arrangement.'

Table 5. Signals of risk selection in the Dutch basic health insurance

Signal	Description
Selection actions intended to break the pooling arrangements	
1	***Risk selection via 'twin products'*** – Some insurers offer nearly-identical basic-insurance products with different prices in combination with different options for supplementary products. In general, the lower-priced twin comes with less generous (supplementary) coverage than the higher-priced twin.
2	***A warning from the largest insurance conglomerate*** – In 2011 the largest health insurer in the Netherlands wrote a letter to the Dutch Parliament to warn the government that the predictable loss on chronically ill people discourages insurers to invest in the quality of care for this group.
3	***Targeting of highly-educated people*** – An insurance product called 'Promovendum' explicitly targeted highly-educated people (at the time this group was known to be predictably profitable to insurers).
4	***Targeting of foreign seasonal workers*** – Some insurers offer(ed) advantageous products to foreign seasonal workers (at the time this group was known to be predictably profitable to insurers).
Selection actions that break the pooling arrangements	
5	***Risk segmentation across insurers*** – Academic research has provided hard evidence of risk segmentation across insurers. During the switching period 2008-2009, most insurers were confronted with a disproportionate inflow and/or disproportionate outflow of either low- or high-risk people.
6	***Risk segmentation across insurance products*** – Several studies, in different time periods, have provided evidence or signals of risk segmentation across insurance products. In general, voluntary deductibles and narrow-network products are found to attract a disproportionate share of low-risk people while more generous products are found to attract a disproportionate share of high-risk people.
7	***Risk segmentation across group arrangements*** – Several studies indicate that premium discounts for specific group arrangements are not driven by cost efficiency, which has resulted in the hypothesis that these discounts are driven by risk segmentation rather than cost efficiency.
8	***Choice of insurance product by pregnant women –*** A disproportionate share of pregnant women chose an insurer that offered a product that was particularly attractive to pregnant women. Because at that time the Dutch risk equalization model lacked risk adjusters that explicitly indicate pregnancy, this insurer was confronted with substantial losses (i.e., thousands of Euros per pregnant woman per year).

Note: a more detailed description of these signals is provided in the Appendix.

4.5. Conclusion

This section has discussed the measures used for ex-ante evaluation of risk equalization and risk sharing in a well-developed policy context: the Dutch health insurance for curative care. In general, we can conclude that the Evaluation Framework used in the official policy cycle is incomplete and partly invalid. First, the framework does not look at the entire spectrum of potential effects mentioned in Section 2.2. For example, none of the potential positive effects

in Table 2 are explicitly considered in the Evaluation Framework, neither qualitatively nor quantitatively. Second, the set of quantitative measures in the Evaluation Framework is very limited. While the 'mean per person financial result for subgroups' is a meaningful indicator of selection incentives for insurers, the set of groups for which this measure is calculated is very limited. Important groups – such as groups with specific diseases that insurers can select against – are missing. Third, the Evaluation Framework heavily relies on measures of statistical fit such as the R-squared and CPM (calculated at the individual level). As discussed in Section 3 of this paper such measures are hardly meaningful for indicating the potential effects of risk equalization and risk sharing. Finally, at the subgroup and insurer-portfolio level, the Evaluation Framework includes six measures that are not considered in academic research (see Section 3). All six measures are hard to interpret and are no valid measures to assess the impact of risk equalization on the incentives for risk selection. Signals of risk selection in the Netherlands underline the relevance of revising and extending the Evaluation Framework.

5. Conclusion and discussion

In this paper we have provided an overview of the potential positive and negative effects of risk equalization and risk sharing in regulated competitive health insurance markets. We have also examined how and to what extent these potential effects are considered in research and policy. Our findings lead us to three main conclusions. First, the potential effects of risk equalization and risk sharing are multidimensional: in total, we distinguished 22 potential effects, most of which relate to efficiency and some (also) to fairness (see Section 2.2.). Of these potential effects, 14 work in a positive direction, either via a reduction of 'actions that intend to break the pooling arrangements' (effects 1-7 in Table 2) or via a reduction of 'the breaking of pooling arrangements' (effects 8-14 in Table 2). Five potential effects work in a negative direction and relate to the direct cost of risk equalization and risk sharing or the indirect costs due to perverse incentives. Three potential effects relate to selection by consumers in/out the market (which are not relevant for markets with an effective mandate) and work in a positive or negative direction, dependent on how risk equalization and risk sharing payments are financed.

Our second conclusion is that the academic literature offers a wide range of quantitative measures for ex-ante evaluation of risk equalization and risk sharing schemes. In the light of the potential effects, however, some measures are more informative than others, which implies that the choice of measures should be made carefully. Moreover, we find that most of the available measures do not go beyond incentives. The development of more sophisticated measures that incorporate the impact of incentives on behavioral actions is an important step to better predict the potential effects of alternative risk equalization/sharing designs.

A third conclusion is that in the Dutch basic health insurance for curative care policymakers do not consider all potential effects. None of the potential positive effects of risk equalization and risk sharing mentioned in Table 2 are explicitly considered in the Evaluation Framework. Moreover, the Evaluation Framework includes a series of inappropriate measures that poorly link to the potential effects, such as the R-squared (i.e., the proportion of variance in spending explained by the payment model) and measures of fit at the level of the risk composition of the insurers' portfolios three years ago. We think this is problematic since an incomplete and/or inappropriate set of evaluation measures can lead to biased conclusions about the

performance of risk equalization and risk sharing models resulting in suboptimal policy making. Signals of risk selection in the Netherlands underline the importance of considering the entire spectrum of potential effects (except for the effects that relate to selection in and out of the market).[7]

Our conclusions lead us to two key recommendations, one for academic researchers and one for policymakers. Our recommendation for academics researchers is to develop additional meaningful measures for predicting the effects of risk equalization and risk sharing. This requires going beyond under/overcompensation by developing economic models of how incentives translate into behavior and effects, which might strongly depend on the institutional context and thus vary across systems. Existing models, such as those developed by Frank et al. (2000) and Einav et al. (2010) provide a strong basis for this line of research. The challenge will be to customize such models to the setting of interest and to build such models for the other potential effects of risk equalization and risk sharing (the models developed by Frank et al. and Einav et al. link to a small subset of the 22 potential effects). We realize that it might be too ambitious to develop a complete model for each potential effect in each context. In many settings and for many effects, empirical estimations on how insurers and consumers respond to incentives might not be available. In these cases, a first important step might be to construct a conceptual 'blueprint' of how incentives translate into potential behavior and effects.

Our recommendation for policymakers is to critically review and revise their evaluation framework in the light of the 22 potential effects identified in this paper. Although our conclusion that the Dutch Evaluation Framework is incomplete only holds for the Netherlands, we expect that evaluation frameworks used in other countries are incomplete too, given that the Netherlands is known to have a relatively well-develop policy and research cycle. As a first step, it is crucial for policymakers to carefully consider which of the 22 effects are relevant for their insurance system. This depends on system features such as the presence of an insurance mandate, the set of choice options for consumers and the flexibility of insurers regarding the design of insurance products. As a second step, it is important to accurately describe the relevant effects in the evaluation framework. As a third step, it is crucial to select meaningful quantitative measures for indicating the relevant effects. Although academic research provides guidance on the choice of quantitative measures, we realize that the choice of measures might be subject to data restrictions. In case there is no data (or no meaningful measure) available to quantify a potential effect, we recommend including that potential effect in a qualitative way, e.g., via discussions with experts and stakeholders. By considering the full spectrum of potential effects, priority can be given to achieving the effects that are socially most desirable.

[7] We did not explicitly examine the occurrence of the negative effects mentioned in Table 3. Until recently, no research existed on this topic for the Netherlands. This year, however, the Dutch 'National Healthcare Institute' quantitatively mapped the impact of risk equalization on incentives for insurers to choose efficient treatments over inefficient treatments. They conclude that in some cases risk equalization reduces or even eliminates these incentives, e.g., when the inefficient treatments trigger a higher risk equalization payment (e.g., via a diagnostic cost group) while the efficient treatments do not. In a qualitative study, the National Healthcare Institute finds signals of insurers being responsive to this type of perverse incentives (ZIN, 2022). We believe that empirical estimations on the occurrence and effects of such behavior is an important direction for future research. Recent international research confirms that perverse incentives from risk equalization can lead to undesirable behavior of health insurers (Geruso and Layton, 2020).

Appendix

Signal 1: Risk selection via 'twin products'

More and more health insurers in the Netherlands offer virtually identical basic-insurance products with large price differences (Elferink, 2021). Although insurers are not allowed to offer *perfectly identical* basic-insurance products with different prices, some of them offer *nearly-identical* basic-insurance products with different prices. The two products are nearly identical except for small differences, e.g., in terms of coverage or options regarding customer service. Simultaneously, these insurers often differentiate the supplementary-insurance options that come with these basic-insurance products: in general, the supplementary-insurance options are much more limited for the lower-priced basic product than for its higher-priced twin. It is hypothesized that health insurers offering such twin products are looking for healthy enrollees that are – on average – overcompensated by the risk equalization system. Once a health insurer starts offering twin products, other insurers must follow this strategy in order to prevent their relatively healthy insured from switching to the lower priced twin product.

Signal 2: A warning from the largest insurance conglomerate

In a letter to the Dutch parliament, the largest health insurer in the Netherlands warns that the quality of care for people with a chronic disease is under pressure since – on average – this group is predictably unprofitable for insurers. The insurer argues that this predictable loss discourages insurers to invest in the quality of care for vulnerable groups and undermines the cross subsidies from healthy people to those with a chronic disease (Achmea, 2011).

Signal 3: Targeting of highly educated people

According to a member of Parliament, in 2012, a specific insurance product called 'Promovendum' was mainly concerned with targetting highly educated people with the aim of avoiding people with a chronic disease (Tweede Kamer, 2014). In addition to selective marketing towards highly educated people, this product offered enrollees the option to cancel their insurance at any time contract during the year. From the perspective of Promovendum, and potentially also from the perspective of enrollees (in particular those who opted for a voluntary deductible and/or narrow-network plan), it is financially beneficial when enrollees switch to another insurer when they become ill. Offering low-priced products to profitable groups like highly educated people, students and young people was called undesirable by a committee that was commissioned by the Minister of Health to assess the performance of the risk equalization scheme in 2012 (Commissie Don, 2012; Van Dorresteijn, 2012; Schippers, 2012).

Signal 4: Targeting of foreign seasonal workers

There has been a parliamentary discussion about a number of health insurers who offer supplementary health insurance products for migrants who come to work in the Netherlands (NZa, 2014). Since these workers are typically relatively healthy (and often return to their home country in case of a serious medical condition, which often implies a termination of their health insurance contract), they are predictably profitable to health insurers (Vermeulen et al., 2021). To attract this group, some insurers offer specific supplementary insurance products that cover the out-of-pocket spending under the deductible in the basic health

insurance. For example, the insurer HollandZorg is a self-acclaimed specialist and market leader in health insurance for migrant workers (HollandZorg, 2019). Via a supplementary insurance (only available for seasonal workers), they cover out-of-pocket spending under both the mandatory deductible and the voluntary deductible in the basic insurance. Since this supplementary insurance has no or a very low premium, this offer is very attractive to seasonal workers.

Signal 5: Risk segmentation across insurers

By looking at people who switched insurer from 2008 to 2009, Van de Ven et al. (2017) have provided the first 'hard' empirical evidence on risk selection in the Dutch basic health insurance market. By isolating selection effects from cost efficiency, this study finds significant market segmentation. The authors argue that the findings are likely to provide an underestimation of the 'true' market segmentation, one of the reasons being that most insurers offer multiple products (some of which potentially attract low risks while others potentially attract high risks).

Signal 6: Risk segmentation across insurance products

Among a total of 74 insurance products in the period 2010-2013, there are strong indications for segmentation of policyholders for 7 products with a selective population and a selective inflow and outflow (NZa, 2016a). For another 20 products the indications for segmentation are smaller as that there is a selective inflow and outflow too, but their overall population is not selective. They also find an indication that segmentation in some cases also takes place among insurers within the same conglomerate. At the conglomerate level, however, the results of insurance products cancel out. Furthermore, it is found that risk equalization results are above average for groups of insured with a high voluntary deductible and groups of insured with a narrow-network plan. At the same time, these insured pay a lower-than-average premium. It is hypothesized that the self-selection of consumers into voluntary deductible and narrow-network plans partly explain the difference in risk equalization results among products.

In a qualitative follow-up research, it was concluded that the health insurance market achieves risk solidarity given that at the conglomerate level insurers strive for a so-called "fair share" (i.e., representative share) of the market. Nevertheless, the health insurance market is segmented at the level of insurers and products (NZa, 2016b). Indirect premium differentiation (i.e., different premiums for nearly identical products offered within a conglomerate) is mentioned as one of the causes for segmentation. In addition, selection actions occur by intermediaries and group arrangements for which controls by the associated insurers are not strict enough.

More recently, the Healthcare Authority concluded that the signals of risk selection have increased in the period 2015-2018 compared to the period 2014-2017: segmentation in the health insurance market has increased, management through marketing seems effective and products with limited contracting for specialist medical care (narrow-network products) have significantly healthier insured populations (NZa, 2020).

Another study, by Hamstra et al. (2022) shows that the composition of products in the years 2016-2019 is very selective with respect to health risk, which has both a positive and negative impact on the financial results of these products. This relationship is somewhat weaker at the

aggregation level of insurers, who typically offer multiple insurance products. The characteristics of the products have a major influence on the emergence of selective portfolios. Large differences in results among products are mainly caused by selective inflows and selective outflows of (un)healthy enrollees. A consumer information surplus about (expected) developments in individual health may be an important explanation for the emergence of selective portfolios. Moreover, most switchers do not switch each year.

Based on public sources containing financial statements of Dutch health insurers about the composition of their premiums in 2019-2022, it was concluded that the existing premium differences among insurers could not be explained by differences in their health care purchasing policies alone. Moreover, it was concluded that the health risk profile of their insured people determined the financial results more than their health care purchasing policy would (Egberts & Hamstra, 2019; Egberts, 2020; Egberts et al., 2021; Egberts et al., 2022).

Signal 7: Market segmentation across group arrangements

For a given insurance product, insurers are allowed to make an arrangement with a group of enrollees, e.g., the employees of a firm or members of a patient organization. The goal of such arrangements (as intended by the regulator) is to improve efficiency. For example, these arrangements allow insurers to pass through some of the cost efficiency gains from managed care for a specific disease to the group of patients with this disease, which promotes enrollment of these patients in efficient plans. Insurers were allowed to provide a premium discount of 10 percent for the basic insurance (reduced to 5 percent in 2020-2022 and to 0 percent in 2023). In addition, discounts can be provided for supplementary insurance products (for which no special regulation exists). While these group arrangements are meant to improve efficiency, they can also lead to risk selection. If different risk types sort into different group arrangements, premium discounts might start to reflect difference in health rather than differences in cost efficiency.

In an evaluation of the functioning of the Dutch Health Insurance Act over the period 2009-2014, it was concluded that descriptions of group arrangements hardly mentioned aspects of efficiency (KPMG, 2014). Therefore, it is expected that in these group arrangements there will be limited attention for actually improving efficiency. It is hypothesized that premium discounts for specific groups contracts are based on selection rather than efficiency. Per 2023, the premium discount for basic coverage is no longer allowed. Nevertheless, insurers will still be able to provide an unlimited group discount on the premium for supplementary insurance.

Signal 8: Choice of insurance product by pregnant women

An above average number of pregnant women chose health insurer Eno since this insurer offered an insurance product that was particularly attractive to pregnant women (who expect to deliver a baby in the contract year). Given that – until 2023 – the risk equalization model lacked risk adjusters that explicitly indicate pregnancy, this insurer was confronted with substantial losses of – on average – thousands of Euros per pregnant woman per year. (NZa, 2016b).

References

ACM/NZA. (2018). Beter kiezen op de polismarkt. (English translation: Making better choices on the health insurance market.) Available at: https://www.acm.nl/sites/default/files/documents/2018-07/acm-nza-rapport-beter-kiezen-op-de-polismarkt.pdf (Accessed: 8 November 2022).
Achmea. (2011). Brief van Achmea aan de leden van de Vaste Commissie voor Volksgezondheid, Welzijn en Sport. Den Haag: Zeist, 7 september. (English translation: Letter from Achmea to the members of the committee for health, welfare and sports.)

Ash, A.S., R.P. Ellis, G. Pope, et al. (2000). Using diagnoses to describe populations and predict costs. Health Care Financing Review, 21: 7-28.

Bauhoff, S. (2012). Do health plans risk select? An audit study on the Germany's social health insurance. Journal of Public Economics, 96: 750-759.

Beaulieu N., D.M. Cutler, K. Ho, G. Isham, T. Lindquist, A. Nelson & P. O'Connor. (2006). The business case for diabetes disease management for managed care organizations. Forum Health Economic Policy, 9:1–36.

Beck K., L. Kauer, T.G. McGuire & C.P.R. Schmid. (2020) Improving risk-equalization in Switzerland: Effects of alternative reform proposals on reallocating public subsidies for hospitals. Health Policy, 124: 1363-1367.

Bundorf, M.K., J. Levin & N. Mahoney. (2012). Pricing and Welfare in Health Plan Choice. American Economic Review, 102: 3214-48.

Cabral M., M. Geruso & N. Mahoney. (2018). Do Larger Health Insurance Subsidies Benefit Patients or Producers? Evidence from Medicare Advantage. American Economic Review, 108: 2048-87.

Cao Z. & T.G. McGuire. (2003). Service-level selection by HMOs in Medicare. Journal of Health Economics, 22: 915-931.

Commissie Don (2012). Evaluatie Risicoverevening Zorgverzekeringswet. (English translation: Evaluation of Risk Equalization under the Health Insurance Act.) Available at: https://www.eumonitor.nl/9353000/1/j4nvgs5kjg27kof_j9vvik7m1c3gyxp/vj0be75kk7zz/f=/blg171977.pdf (Accessed: 8 November 2022).

Duijmelinck, D. et al. (2013) Overstapgedrag en risicoselectie op de zorgverzekeringsmarkt. iBMG. (English translation: Switching and risk selection in the health insurance market.) Available at: https://puc.overheid.nl/nza/doc/PUC_2972_22/ (Accessed: 8 November 2022).

Egberts, B.W. & G. Hamstra (2019). Premieverschillen tussen zorgverzekeraars nemen toe. Equalis Strategy & Modeling. (English translation: Premium variation across health insurers is increasing.) https://equalis.nl/premieverschillen-tussen-zorgverzekeraars-nemen-toe/.

Egberts, B.W. (2020). Verschillen in het resultaat van verevening en zorginkoop tussen verzekeraars blijven groot. (English translation: Variation in financial results from risk equalization and healthcare purchasing across insurers continue to be substantial.) Equalis

Strategy & Modeling. https://equalis.nl/verschillen-in-het-resultaat-van-verevening-en-zorginkoop-tussen-verzekeraars-blijven-groot/.

Egberts, B.W., G. Hamstra & R. van Ingen Schenau (2021). Hogere kosten Covid-19 zetten de zorgpremies van 2021 onder druk. Equalis Strategy & Modeling. (English translation: Higher spending due to Covid-19 puts pressure on insurance premiums for 2021.) https://equalis.nl/hogere-kosten-covid-19-zetten-de-zorgpremies-van-2021-onder-druk/.

Egberts, B.W., G. Hamstra & L. Tjon Joe Gin (2022). Flinke rem op premiegroei door grootse inzet van reserves. (English translation: Insurers use their financial reserves to slow down the increase in health insurance premiums.) Equalis Strategy & Modeling. https://equalis.nl/flinke-rem-op-premiegroei-door-grootse-inzet-van-reserves/.

Einav, L. & A. Finkelstein. (2011). Selection in Insurance Markets: Theory and Empirics in Pictures. Journal of Economic Perspectives, 25: 115-38.

Einav L., A. Finkelstein & M.R. Cullen. (2010). Estimating Welfare in Insurance Markets Using Variation in Prices. The Quarterly Journal of Economics, 125: 877–921,

Elferink, T. (2021) Zorgverzekeraars lokken gezonde klanten met goedkope tweelingpolissen, NU. (English translation: Health insurers are targeting healthy consumers by means of cheap 'twin products'.) Available at: https://www.nu.nl/gezondheid/6175550/zorgverzekeraars-lokken-gezonde-klanten-met-goedkope-tweelingpolissen.html (Accessed: 29 March 2022).

Ellis, R.P., H.E. Hsu, J.J. Siracuse, A.J. Walkey, K.E. Lasser, B.C. Jacobson, C. Andriola, A. Hoagland, Y. Liu, C. Song, T.C. Kuo & A.S. Ash. (2022). Development and Assessment of a New Framework for Disease Surveillance, Prediction, and Risk Adjustment: The Diagnostic Items Classification System. JAMA Health Forum 3: e220276.

Ellis R.P., S. Jiang & T.C. Kuo. (2013). Does service-level spending show evidence of selection across health plan types? Applied Economics, 45: 1701-1712

Ellis R.P, B. Martins & S. Rose. (2018). Risk Adjustment for Health Plan Payment. In McGuire, T.G. & van Kleef, R.C. (Eds.). Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets, Elsevier Publishing.

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

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

Geruso, M. & Layton, T. (2020). Upcoding: Evidence from Medicare on Squishy Risk Adjustment. Journal of Political Economy, 128: 984-1026.

Geruso, M., T.J. Layton, G. McCormack & M. Shepard. (2021). The Two Margin Problem in Insurance Markets. The Review of Economics and Statistics. doi: https://doi.org/10.1162/rest_a_01070

Geruso, M., T. Layton, & D. Prinz. (2019). Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges. American Economic Journal: Economic Policy, 11: 64-107.

Geruso, M., & T.G. McGuire. (2016). Tradeoffs in the design of health plan payment systems: Fit, power and balance. Journal of Health Economics. 47: 1-19.

Glazer, J. & T.G. McGuire. (2002). Setting health plan premiums to ensure efficient quality in health care: minimum variance optimal risk adjustment. Journal of Public Economics, 84: 153-173.

Hackmann, M.B., J.T. Kolstad, & A.E. Kowalski. (2015). Adverse Selection and an Individual Mandate: When Theory Meets Practice. The American Economic Review, 105: 1030–1066.

Hamstra, G., P. Van Drunen, R. Hoekstra, M. Mol. & P. Stam. (2022) Restrisico's 2: Vereveningsresultaat van verzekeraars en polissen WOR 1074 (English translation: Remaining problems in risk equalization: Under/overcompensation of insurers and insurance products.) Equalis - Strategy & Modeling. Available at: https://www.rijksoverheid.nl/documenten/rapporten/2022/01/26/restrisicos-2-vereveningsresultaat-van-verzekeraars-en-polissen (Accessed: 17 March 2022).

Han, T., & K. Lavetti. (2017). Does Part D Abet Advantageous Selection in Medicare Advantage? Journal of Health Economics, 56: 368-382.

Handel, B.R. & J.T. Kolstad. (2015). Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare. American Economic Review, 105: 2449-2500.

Kanters, T., W.B.F. Brouwer, R.C.J.A. van Vliet, P.H.M. van Baal & J.J. Polder. (2013). A new prevention paradox: The trade-off between reducing incentives for risk selection and increasing the incentives for prevention for health insurers. Social Science & Medicine, 76: 150-158.

Kong, E., T.J. Layton & M. Shepard. (2022). Adverse Selection and Insurer Participation in Health Insurance Markets. Working paper. https://scholar.harvard.edu/files/mshepard/files/adverse_selection_pricing_LaytonKongShepard.pdf

KPMG. (2014) Evaluatie Zorgverzekeringswet. Eindrapportage. (English translation: Evaluation of the Health Insurance. Final report.) Act Available at: https://zoek.officielebekendmakingen.nl/blg-384793.pdf (Accessed: 8 November 2022).

HollandZorg. (2019). Folder HollandZorg 2020. (English translation: Brochure HollandZorg 2020). Available at: https://www.werkpluszorg.nl/wp-content/uploads/2019/12/Folder-HollandZorg-2020.pdf (Accessed: 27 June 2022).

Lamers, L.M. (2001). Health-Based Risk Adjustment: Is Inpatient and Outpatient Diagnostic Information Sufficient? Inquiry, 38: 423–431.

Lavetti, K., & K. Simon. (2018). Strategic Formulary Design in Medicare Part D Plans. American Economic Journal: Economic Policy, 10: 154-92.

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

Layton, T.J., T.G. McGuire & R.C. van Kleef. (2018). Deriving Risk Adjustment Weights to Maximize Efficiency of Health Insurance Markets. Journal of Health Economics, 61: 93-110.

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

McGuire, T.G., S. Schillo & R.C. van Kleef. (2020). Reinsurance, Repayments, and Risk Adjustment in Individual Health Insurance: Germany, The Netherlands and the U.S. Marketplaces. American Journal of Health Economics, 6: 139–168.

McGuire, T.G., S. Schillo & R.C. van Kleef. (2021). Very High and Low Residual Spenders in Private Health Insurance Markets: Germany, The Netherlands and the US Marketplaces. European Journal of Health Economics, 22: 35–50.

McGuire, T.G. & R.C. van Kleef. (2018). Risk Sharing. In: McGuire, T.G. & van Kleef, R.C. (Eds.), Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets, Elsevier Publishing.

Montz, E., T. Layton, A.B. Busch, R.P. Ellis, S. Rose & T.G. McGuire. (2016). Risk-Adjustment Simulation: Plans May Have Incentives To Distort Mental Health And Substance Use Coverage. Health Affairs, 35: 1022-1028.

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

Newhouse J.P., M. Price, J.M. McWilliams, J. Hsu, T.G. McGuire. (2015). How much favorable selection is left in Medicare Advantage? American Journal of Health Economics, 1: 1-26.

Nyman, J.A. (1999). The economics of moral hazard revisited. Journal of Health Economics, 18: 811-824,

NZa. (2016a). Kwantitatief onderzoek naar risicoselectie en risicosolidariteit op de zorgverzekeringsmarkt. (English translation: Quantitative research on risk selection and risk solidarity in the Dutch health insurance market.) Available at: https://zoek.officielebekendmakingen.nl/blg-672551.pdf. (Accessed: November 9, 2022).

NZa. (2016b). Risicoselectie en risicosolidariteit zorgverzekeringsmarkt: Kwalitatief onderzoek 2016. (English translation: Risk selection and risk solidarity in the Dutch health insurance market: Qualitative research 2016.) NZa. Available at: https://puc.overheid.nl/nza/doc/PUC_3481_22/1/. (Accessed: November 9, 2022).

NZa. (2020). Monitor Zorgverzekeringen 2020. (English translation: Monitor of the basic health insurance for curative care 2020). Available at: https://www.rijksoverheid.nl/documenten/rapporten/2020/09/10/monitor-zorgverzekeringen-2020. (Accessed: November 9, 2022).

Pope, G.C., J. Kautter, R.P. Ellis, A.S. Ash, J.Z. Ayanian, L.I. Iezzoni, J.L. Melvin, M. Jesse & J. Robst. (2004). Risk adjustment of Medicare capitation payments using the CMS-HCC model. Health Care Financing Review, 25: 119-141.

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

Saltzman, E. (2021). Managing Adverse Selection: Underinsurance vs. Underenrollment." The RAND Journal of Economics, 52: 359-381.

Schippers, E.I. (2012) Herziening Zorgstelsel; Brief regering; Commissie evaluatie risicoverevening ZVW. (English translation: Healthcare system reform; Letter by the government; Committee for evaluation of risk equalization.) Available at: https://zoek.officielebekendmakingen.nl/kst-29689-405.html (Accessed: 29 June 2022).

Schmid C.P. & K. Beck. (2016). Re-insurance in the Swiss health insurance market: Fit, power, and balance. Health Policy, 120: 848-55.

Shepard, M. (2022). Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange. American Economic Review, 112: 578-615.

Stam P.J.A., R.C.J.A. van Vliet & W.P.M.M. van de Ven. (2010a). A limited-sample benchmark approach to assess and improve the performance of risk equalization models. Journal of Health Economics, 29: 426-37.

Stam, P.J.A., R.C.J.A. van Vliet & W.P.M.M. van de Ven. (2010b). Diagnostic, Pharmacy-Based, and Self-Reported Health Measures in Risk Equalization Models. Medical Care, 48: 448–57.

Trottmann M., P. Zweifel & K. Beck. (2012). Supply-side and demand-side cost sharing in deregulated social health insurance: which is more effective? Journal of Health Economics, 31: 231-42.

Tweede Kamer. (2014). Risicoselectie in de basisverzekering. (English translation: Risk selection in the basic health insurance.) Available at: https://www.tweedekamer.nl/debat_en_vergadering/plenaire_vergaderingen/details/activiteit?id=2014A00090. (Accessed: November 9, 2022).

Van Barneveld, E.M., L.M. Lamers, R.C.J.A. van Vliet & W.P.M.M. van de Ven. (2001). Risk sharing as a supplement to imperfect capitation: a tradeoff between selection and efficiency. Journal of Health Economics, 20: 147-168.

Van Barneveld, E.M., Lamers, L.M., Van Vliet, R.C.J.A. et al. (2000). Ignoring small predictable profits and losses: a new approach for measuring incentives for cream skimming. Health Care Management Science, 3: 131–140.

Van de Ven, W.P.M.M., K. Beck, F. Buchner, E. Schokkaert, F.T. Schut, A. Shmueli et al. (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 de Ven, W.P.M.M. & R.P. Ellis. (2000). Risk adjustment in competitive health insurance markets. In: Culyer AJ, Newhouse JP, editors. Handbook of Health Economics (Chapter 14) Amsterdam: Elsevier, 755-845.

Van de Ven, W.P.M.M., G. Hamstra, R. van Kleef, M. Reuser & P. Stam (2022). The goal of risk equalization in regulated competitive health insurance markets. European Journal of Health Economics: 10.1007/s10198-022-01457-7

Van de Ven, W.P.M.M., R.C. van Kleef & R.C.J.A. van Vliet. (2017). How can the regulator show evidence of (no) risk selection in health insurance markets? Conceptual framework and empirical evidence. The European Journal of Health Economics, 18: 167-180.

Van Dorresteijn, M. (2012). Commissie waarschuwt voor risicoselectie zorgverzekeraars. (English translation: Committee warns for risk selection by health insurers.) Zorgvisie, 8 June 2012. Available at: https://www.zorgvisie.nl/commissie-waarschuwt-voor-risicoselectie-zorgverzekeraars-zvs014109w/ (Accessed: 29 June 2022).

Van Kleef, R.C., F. Eijkenaar & R.C.J.A. van Vliet (2019). Selection Incentives for Health Insurers in the Presence of Sophisticated Risk Adjustment. Medical Care Research and Review, 77: 584-595.

Van Kleef, R.C., F. Eijkenaar, R.C.J.A. van Vliet & M. Nielen. (2020). Exploiting Incomplete Information in Risk Adjustment Using Constrained Regression. American Journal of Health Economics, 6: 477–497

Van Kleef, R.C., F. Eijkenaar, R.C.J.A. van Vliet & W.P.M.M. van de Ven. (2018b). Health Plan Payment in the Netherlands. In McGuire, T.G. & van Kleef, R.C. (Eds.), Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets, Elsevier Publishing.

Van Kleef, R.C., F.T. Schut & W. P. M. M. van de Ven. (2018a). Premium Regulation, Risk Equalization, Risk Sharing, and Subsidies: Effects on Affordability and Efficiency. In McGuire T.G and R.C. van Kleef (eds). Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice, Elsevier

Van Kleef, R.C. & R.C.J.A. van Vliet. (2011). Prior Use of Durable Medical Equipment as a Risk Adjuster for Health-Based Capitation, Inquiry, 47: 343-358.

Van Kleef, R.C. & R.C.J.A. van Vliet. (2022) How to deal with persistently low/high spenders in health plan payment systems? Health Economics, https://doi.org/10.1002/hec.4477

Van Veen, S.C.H.M., R.C. van Kleef, W.P.M.M. van de Ven & R.C.J.A. van Vliet. (2015). Is There One Measure-of-fit that Fits All? A Taxonomy and Critical Assessment of Measures

that are used for assessing the Predictive Performance of Risk-Equalization Models. Medical Care Research and Review, 72: 220–243.

Vermeulen, W., L. Kroon & C. Odding. (2021). Buitenlandse seizoenarbeiders in de risicoverevening. (English translation: Foreign seasonal workers in risk equalization.) SEO Economisch Onderzoek, p. 66. Available at: https://www.rijksoverheid.nl/documenten/rapporten/2021/11/30/buitenlandse-seizoenarbeiders-in-de-risicoverevening.

VVD & PVDA. (2012) Regeerakkoord 2012: Bruggen slaan. (English translation: Coalition agreement 2012: Bridging gaps.) Available at: https://www.rijksoverheid.nl/documenten/rapporten/2012/10/29/regeerakkoord (Accessed: 8 November 2022).

VVD, CDA, D66 & CU. (2017). Regeerakkoord 2017: Vertrouwen in de toekomst. (English translation: Coalition agreement 2017: Trust in the future.) Available at: https://www.rijksoverheid.nl/documenten/publicaties/2017/10/10/regeerakkoord-2017-vertrouwen-in-de-toekomst (Accessed: 8 November 2022).

VWS. (2016). Position papers ten behoeve van symposium: Toekomst van de risicoverevening. (English translation: Position papers presented at the symposium on the future of risk equalization.) Available at: https://www.tweedekamer.nl/downloads/document?id=2016D25308 (Accessed: 8 November 2022).

Withagen-Koster A.A., R.C. van Kleef & F. Eijkenaar. (2018). Examining unpriced risk heterogeneity in the Dutch health insurance market. European Journal of Health Economics, 19: 1351-1363.

Withagen-Koster, A.A., R.C. van Kleef & F. Eijkenaar. (2020). Incorporating self-reported health measures in risk equalization through constrained regression. The European Journal of Health Economics, 21: 513–528

Withagen-Koster, A.A., R.C. van Kleef & F. Eijkenaar. (2022). An innovative method to mitigate risk selection incentives in health insurance markets with sophisticated risk equalization: High-risk pooling based on health survey information. Under review.

WOR 871. (2017). Toetsingskader Risicoverevening. (Evaluation Framework for Risk Equalization).

ZIN (2022), Advies Passende Zorg en Risicoverevening. (English translation: Recommendation on prudent care and risk equalization) https://www.zorginstituutnederland.nl/binaries/zinl/documenten/adviezen/2022/08/16/advies-passende-zorg-en-risicoverevening/Rapportage+onderzoek+Risicoverevening+%26+Passende+Zorg.pdf

Zink, A. & S. Rose. (2021) Identifying undercompensated groups defined by multiple attributes in risk adjustment. BMJ Health Care Inform 28: e100414. doi: 10.1136/bmjhci-2021-100414.

Erasmus University Rotterdam
Erasmus Centre for Health Economics Rotterdam
Burgemeester Oudlaan 50
3062 PA Rotterdam, The Netherlands
T +31 10 408 8555
E escher@eur.nl
W www.eur.nl/escher