

UC Santa Barbara

Departmental Working Papers

Title

Why 'Optimal' Payment for Healthcare Providers Can Never Be Optimal Under Community Rating

Permalink

<https://escholarship.org/uc/item/37b9q0k3>

Authors

Frech, Ted E
Zweifel, Peter

Publication Date

2015-04-27

Why ‘Optimal’ Payment for Healthcare Providers Can Never Be Optimal Under Community Rating

Peter Zweifel and H.E. Frech III

This version: 10 Apr 2015

Abstract

This article extends the received literature on optimal provider payment by accounting for consumer heterogeneity in preferences for health insurance and health care. This heterogeneity breaks down the separation of the relationship between providers and the health insurer and the relationship between consumers and the insurer. Both experimental and market evidence for a high degree of heterogeneity are presented. Given heterogeneity, a uniform policy fails to effectively control moral hazard, while incentives for risk selection created by community rating cannot be neutralized through risk adjustment. Consumer heterogeneity spills over into relationships with providers, such that a uniform contract with providers cannot be optimal either. The decisive condition for ensuring optimality of provider payment is to replace community rating (which violates the principle of marginal cost pricing) by risk-rating of contributions combined with subsidization targeted at high risks with low incomes.

Keywords

Community Rating, Health Insurance, Optimal Payment

JEL codes

C93, D82, I13, P22

Introduction

Paying for healthcare providers faces the dual challenge of moral hazard and selection: If reimbursed according to actual expenditure, providers lack incentives to control cost; if made to bear financial risk in the guise of prospective payment, they may try to reduce cost by lowering the quantity or quality of treatment, engage in patient selection, and opt out of the contract. Ever since Chalkley and Malcomson (1998a, 1998b), the dominant approach has been to determine a share of healthcare expenditure (HCE) to be borne by providers which balances their moral hazard effects against selection incentives. The suggestion is to make this solution the basis of nationwide prospective payment of healthcare providers, reminiscent of Lerner's (1934, 1936) market socialism.

However, moral hazard effects also characterize patients (Zweifel and Manning, 2000), while risk selection is also a consumer activity because in the presence of imperfect risk adjustment (RA), low risks seek out an insurer who does not burden them with excessive payment into the RA scheme whereas high risks seek out one who promises them a high subsidy thanks to the contribution from RA (Zweifel, 2013a). Therefore, the traditional separation of optimal provider payment from the structuring of insurance contracts is artificial. The present article argues that it is misguided, too because it neglects the fact that community rating (and other forms of premium regulation) prevent an optimal structuring of health insurance contracts, which in turn makes optimal provider payment impossible. More generally, optimal provider payment cannot be attained as long as health insurers are not permitted to tailor their policies to heterogeneous preferences of consumers.

A simple analogy may illustrate the point. Consider a department store, which purchases goods and services from a set of suppliers in the same way as a health insurer does. Both act in anticipation of customers who will exercise options over the products available. Let department store A cater to upscale consumers, with 'Only the best is good enough for you' being its slogan, while department store B serves a more price-conscious clientele. In addition, let there be an unexpected surge in demand. It is evident that A will refrain from pushing its suppliers too hard to increase deliveries in fear of jeopardizing quality, while B may even threaten to change suppliers if the incumbents cannot deliver. Likewise, healthcare financiers, acting on behalf of a clientele with heterogeneous preferences concerning health care, need to negotiate with different types of providers in different ways rather than trying to impose a uniform payment scheme.

The plan of this paper is as follows. Section 2 presents a justification of the argument proffered in the Introduction section, stating that risk selection characterizes not only the behavior of insurers but also consumers. In Section 3, empirical evidence suggesting substantial consumer preference heterogeneity is presented. This evidence is used in Section 4 for a critical review of optimal provider payment as developed by Chalkley and Malcomson (1998a, 1998b). It will be shown that this approach amounts to a simplified form of market socialism in the Lange-Lerner tradition. However, by ignoring preference heterogeneity, it leads to mispricing and inefficiency. Section 5 contains a summary and suggestions designed to permit the development of truly optimal provider payment.

2 The impossibility of shoring up community rating through risk adjustment

Any textbook will state the 'price equal marginal cost' rule early on, emphasizing the inefficiencies that are caused by violations of this rule. And in the event that some consumers cannot pay the market price, the recommended remedy is, in keeping with the second theorem of Welfare Economics, 'Do not regulate price but let the market do its work; simply have the government pay a subsidy to those consumers who cannot afford the good' (Jehle and Reny, 2001, ch. 5.4). Yet in health economics, this advice is often neglected. It would call for risk-rated contributions in health insurance because the (present value of) future expected healthcare expenditure (HCE) constitutes the major part of marginal cost associated with enrolling an extra individual. Rather, much of the literature (with the notable exceptions of Pauly (1970) and Bhattacharya et. al. (2014)) has accepted community rating as a given, although it has the undesirable side effect of inducing inefficient risk selection and incorrect choices. Since an insured population always contains high risks whose (present value of) HCE exceeds the (present value) of contributions under community rating, a health insurer who wants to break even needs to enrol low risks for budget balance.

The regulatory response has been to introduce RA, which serves to artificially increase the marginal cost of a low risk while lowering that of a high risk to the insurer. But this approach has serious problems, ignoring the interaction of selection and more hazard and consumer selection. On the insurer side alone, it proves exceedingly difficult to neutralize incentives for risk selection through RA (Zweifel and Breuer, 2006). For instance, the regulator would have to also know the insurer's planning horizon for applying the relevant rate of discounting as well as the probabilities of transition between risk categories, which are substantial over a few years (Beck et al., 2010). Allowing for moral hazard leads to further problems, discussed below. However, consumers' incentives for risk selection must be neutralized, too. Like an indirect tax, RA payments are ultimately borne by the (low-risk) insured, while high risks receive an implicit subsidy in the guise of the contribution the insurer gets from the RA scheme. Since RA is necessarily highly imperfect, this creates an incentive for the low risks to seek out a health insurer helping them avoid this indirect tax at least in part. Conversely, high risks select an insurer who maximizes the implicit subsidy they stand to obtain. Recent research shows that adverse selection by consumers led to

substantial declines in the purchase of insurance in the U.S. states that mandated it in the 1990s (Clemens 2015). Application of Tinbergen's (1954) rule makes it clear that the single instrument, 'payment into and out of the RA scheme' cannot achieve the two objectives of neutralizing both insurers' and consumers' interest in risk selection (Zweifel, 2013a). Discrepancies between price and marginal cost with all its consequences in terms of inefficiency are therefore certain to be permanent in health insurance subject to community rating (and more generally, premium regulation).

3 Preference heterogeneity regarding health insurance

As argued recently by Martin Feldstein, health economists have too often ignored differences in consumer preferences (2013). The inefficiencies caused by community rating are seriously compounded when consumers exhibit preference heterogeneity with regard to health insurance. There are two types of problems.

(A) Inability of the insurer to control moral hazard in a targeted way

The optimal rate of copayment in the presence of ex-post moral hazard can be shown to depend on no less than five individual-specific parameters (Zweifel et al., 2009, ch. 6.4.2.3):

1. The probability with which the several health states will occur;
2. The level of health associated with each of these states;
3. The marginal utility of consumption associated with health states, which in turn depends on the individual's degree of risk aversion;
4. The elasticity of health w.r.t. the rate of copayment;
5. The optimal level of health as deemed by the individual, given the tradeoffs with other goods.

For copayment to be accepted voluntarily by a risk-averse individual, it must come with a reduction in contribution from the cost of full coverage, the amount of which depends on his or her degree of risk aversion. However, this is not compatible with community rating which in principle requires equality of contributions (or at the very least, equality given a certain rate of copayment). Community rating therefore prevents health insurers from tailoring their contracts to the individual characteristics determining moral hazard.

(B) Reduced insurer incentive to invest in product innovation

In the presence of preference heterogeneity, product innovation serves to match goods and services offered ever more closely to consumer preferences (Lancaster, 1966). In the case of health insurance, premium regulation creates strong incentives for insurers to attract favorable risks through attributes of the contract. This in turn induces regulation designed to prevent it. For instance, exempting pediatric counselling from copayment is likely to attract young families (yet might be justified as a preventive measure in the interest of public health).

More generally, product innovations are first tried by young consumers, likely because finding out about them constitutes an investment (of time at a minimum). For young consumers, the payback period in terms of utility gains is long enough to justify the investment; for old ones, the investment is less likely to pay off. By the same token, product innovation in health insurance usually attracts more young individuals than old ones (see Section 3 for empirical evidence). Even if cream skimming was not the intention, the innovative insurer is suspected of it because its insured population will exhibit a comparatively high share of young individuals. In the presence of RA, this entails a financial sanction because RA formulas invariably include age structure as an adjuster.

One might think that RA at least has the benefit of making the regulator more relaxed about product innovation. Yet RA can have unintended side effects, as evidenced by a seemingly innocuous planned adjustment of the Swiss RA formula. The new formula was to include the dummy variable, 'hospitalization during the previous

year' as an additional adjuster to overcome the poor predictive performance of RA models and the lack of diagnostic information (Swiss healthcare providers have been successfully blocking the transfer of diagnostic data to insurers). A particular social health insurer, who had been a zero net contributor to the RA scheme on average, commissioned an investigation into the financial consequences of this planned change (Schoder et al., 2010). The new formula would have caused it to suddenly pay up to 13 percent of its premium volume into the RA scheme. Since as a social health insurer, it is not legally permitted to hold large reserves, this minor change of the RA formula would have likely caused its insolvency.

An in-depth analysis revealed that the cause was not risk selection; to the contrary, this particular health insurer exhibited an above-average share of enrollees beyond the age of 75. Rather, it had made efforts (hailed by Swiss policy makers keen to achieve the cost savings promised by Managed Care) to keep patients out of hospital. The insurer had built an informal second-opinion network that general practitioners could access before referring a patient to the hospital (and transferring control to a hospital-based specialist in most cases). Although not buttressed by particular financial incentives, this network had proved reasonably effective. Yet the consequence of this laudable initiative was that under the new RA formula, the insurer would have exhibited 'too few' hospitalizations, punishing it for its innovative efforts (the new formula became in effective in 2013 only). Experiences of this type are sure to discourage insurer investment in product innovation and efforts aiming at cost control more generally. Indeed Schoder et. al. (2010) have shown that RA is either necessarily very ineffective or it dramatically undermines cost control incentives or both. Without data on past utilization, RA can adjust for only a very small part of the forward-looking risk. If RA includes past utilization, as shown in the example above, it discourages insurers from controlling costs.

3 Empirical evidence on preference heterogeneity regarding health insurance

Preferably, empirical evidence on preference heterogeneity should be in terms of market outcomes. In the presence of differentiated products, relative marginal utilities associated with product attributes equal relative shadow prices of attributes reflecting not only goods prices but the 'productivity' of goods in terms of attributes (Lancaster, 1966). However, in the case of health insurance, there are distortions at three levels. First, insurance premiums are regulated rather than the outcome of market processes. Second, the prices of goods and especially services entering HCE covered by health insurers are not market prices either (at least outside the United States, where insurers negotiate fees with physicians and hospitals). Third, the quantities entering HCE are influenced by ex-post moral hazard. In this situation, experimental evidence may serve as second-best evidence.

One such source is a Discrete Choice Experiment (DCE) performed in Switzerland, involving some 1,000 participants (Zweifel et al., 2006). Fee-for-service combined with unrestricted physician choice described the status quo in 2003, a contract with several Managed Care (MC) attributes, the alternative. The objective was to measure (negative) willingness-to-pay values for these MC attributes. For instance, Table 1 shows that having competitive social health insurers draw up a physician list based on cost only is resisted so strongly that premiums would have to be 36 percent lower on average to make this restriction acceptable. Regulation in effect at the time limited the premium reduction for MC plans to 20 percent, causing only about 10 percent of the insured (those with a low amount of compensation required) to opt for such a harsh HMO-type variant of MC (for more evidence on preference heterogeneity, see below). However, a physician list taking into account both cost and quality would be accepted in return for 14 percent less premium. Interestingly, choices of respondents revealed that although regional medical centers would promise better quality of treatment thanks to higher surgical volumes, small community hospitals continue to be favored. Finally, mandatory long-term care insurance (the only attribute not related to MC) is also resisted by the young (not shown), even though it would have been financed by a premium surcharge of CHF 50 per month (17 percent of nationwide average premium at the time), to be paid by those aged over 50.

**Table 1. Average WTP values (+: willingness to pay, -: compensation required),
Switzerland (2003)**

Attribute	CHF/month^a	s.e.	%Prem.^b
Physician list based on cost criteria only	-103	13.2	36
Physician list based on quality criteria only	-53	8.8	18
Physician list based on cost and quality criteria	-42	7.8	14
Access to medical innovation delayed 2 years	-65	7.9	22
Reimbursement of generics only	-3	5.5	1
No drugs for minor complaints reimbursed	+6	5.3	2
Choice of hospital restricted to regional centers	-37	5.7	13
Mandatory long-term care insurance	-25	5.8	9
Status quo preference, status quo bias	-59	11.9	20

^a 1 CHF = 1.33 USD at 2003 exchange rates, = 1.08 USD at 2014 exchange rates

^b in percent of the nationwide 2003 premium of CHF 290/month

The bottom line of Table 1 refers to respondents' preference for the status quo. On average, they would have to be compensated by no less than CHF 59 (about 20 percent of nationwide premium, respectively) to move away from the status quo (which was predominantly fee-for-service at the time). Since then, the market share of MC (mostly its 'mild' variants such as IPA-type physician networks) had been continuously increasing, reaching about 50 percent by 2010. This was not deemed sufficient progress by federal politicians, who passed a bill in 2011 calling for MC rather than fee-for-service to become the standard policy in social health insurance. The bill was challenged by a popular referendum and was defeated at the polls by a two-third ma-

majority of voters in June 2012. The voters' verdict suggests that the figures shown in Table 2 are indeed informative of Swiss preferences (Zweifel, 2013b).

**Table 2. Heterogeneity of preferences, WTP values by age group
(CHF, 2003, standard errors in parentheses)**

Attribute	Avge	Age	Age	Age
		25-39	40-64	65 ⁺
Physician list based on cost criteria only ^a	-103 (13.2)	-81 (11.7)	-136 (35.2)	-153 (85.8)
Physician list based on quality criteria only ^b	-53 (8.8)	-33 (7.8)	-72 (22.5)	-133 (77.2)
Physician list based on cost & quality ^b	-42 (7.8)	-29 (7.4)	-60 (19.8)	-76 (49.6)
Access to medical innovation delayed 2 years ^c	-65 (7.9)	-45 (6.7)	-101 (24.5)	-83 (45.6)
Reimbursement of generics only ^c	-3 (5.5)	-9 (5.9)	+4 (11.9)	+24 (27.8)
No drugs for minor complaints reimbursed ^d	+6 (5.3)	+2 (5.6)	+14 (11.9)	+19 (26.9)
Choice of hospital restricted to regional centers ^d	-37 (5.7)	-33 (5.8)	-46 (13.7)	-36 (26.3)

^a The differences between Age 25-39 and Age 40-65 as well as between Age 25-39 and Age 65⁺ are significant at the 0.05 level

^b The difference between Age 25-39 and Age 65⁺ is significant at the 0.05 level

^c The difference between Age 25-39 and Age 40-64 is significant at the 0.05 level

^d None of the differences are significant at the 0.1 level

Heterogeneity of preferences between (and within) age groups becomes apparent in Table 2 (mandatory long-term care insurance is not investigated further). As could be expected, the top age group (65+) needs to be compensated most highly for giving up free physician choice. Additional heterogeneity between regions, income groups, and persons with differing health status is evidenced in Zweifel et al. (2006). Similar DCEs performed in Germany and the Netherlands also show preference heterogeneity between the subjectively healthy and the chronically ill, with the chronically ill resisting MC-type attributes more strongly, possibly because they fear the lock-in effect of being assigned to a gatekeeping physician (MacNeil Vroomen and Zweifel, 2011).

However, between-group heterogeneity is only remotely related to the five individual determinants of moral hazard listed in section 2 above. In Table 2, within-group heterogeneity actually *exceeds* across-group heterogeneity in some instances. For example, WTP values in the 65+ group for the attribute, 'No drugs for minor complaints reimbursed' have such a high standard error (of 26.9) that differences within this group exceed the maximum difference between groups (which amounts to $17 = 29 - 2$ CHF). . Since the comparisons in Table 2 are univariate rather than reflecting predicted values derived from a multivariate probit regression, part of this heterogeneity could be related to other observables such as region, gender, and income group. Yet, substantial within-group heterogeneity caused by unobservable characteristics is likely to remain important, as suggested by the systematic differences between the subjectively healthy and chronically ill found by MacNeil Vroomen and Zweifel (2011).

A particular source of heterogeneity is status quo preference. On the one hand, it reflects risk aversion because the alternative is usually associated with more uncertainty than the well-known status quo. On the other hand, it depends on the probability of reaping the returns on the investment (of time at the least) required to understand an alternative to the status quo. In the case of MC-type health insurance, this entails finding out whether e.g. the physician one is familiar with participates in the network. Table 3 indeed reveals a great deal of heterogeneity in this regard both among participants in the Dutch DCE and the German DCE, respectively (for details, see Leukert-Becker and Zweifel, 2014). In the Netherlands, status quo preference does not unambiguously increase with age, whereas in Germany, it clearly does (as

one would expect based on the investment argument; the difference may also be due to the fact that the top third of the sample starts at age 59 in Germany rather than at age 55 as in the Netherlands). The retired exhibit a stronger status quo preference than the non-retired in both countries, quite likely reflecting an increase in risk aversion after retirement (Halek and Eisenhauer, 2001).

Table 3. Heterogeneity of status quo preferences (WTP values in Euro; Netherlands 2006, Germany 2005)

	Nether-lands	s.e.	p-value χ^2	Germany	s.e.	p-value χ^2
Age < 41 (<43) ^a	-162	35.6	0.06; 5.81	-329	50.1	0.009; 27.7
Age 40-64 (43-59) ^a	-234	42.6		-407	59.0	
Age >55 (>59) ^a	-479	71.0		-940	106.9	
Retired	-456	85.8	0.009; 6.84	-953	119.0	0.000; 19.5
Non-retired	-221	26.9		-402	38.1	
No physician visit ^b	-204	37.4	0.0722; 3.23	-402	68.6	0.122; 2.39
Physician visit(s) ^b	-297	36.1		-533	42.9	
Healthy (subjective)	-164	34.0	0.0017; 9.88	-297	54.1	0.000; 18.95
Ill (subjective)	-325	38.3		-609	49.0	

^a In parentheses: Germany; age groups contain approximately 1/3 of sample

^b During past 12 months

Risk aversion may also be related to health status. An observable indicator is whether or not a person has seen a physician during the past 12 months. In the two countries, compensation required to overcome status quo preference is up to 45 percent higher among respondents with physician visits than among those without any during the past 12 months. However, this between-group heterogeneity is dwarfed by unobservable (to the health insurer at least) within-group differences. Respondents

who are subjectively ill would have to be compensated twice as highly than the subjectively healthy top overcome their status quo preference Findings of this type are relevant for risk selection because unobserved differences in status quo preference may result in sorting processes that cannot be controlled by RA.

Finally, the stage individuals find themselves in a reform process may matter, too. It is remarkable that compensation required to overcome status quo preference is lower in the Netherlands than in Germany throughout. This may be the consequence of the fact that as part of the 2006 reform (which came to an end just prior to the fielding of the DCE), the Dutch were legally obliged to explicitly choose an insurance policy; simply continuing with the existing one was disallowed. They therefore were made to bear the cost of learning to choose, similar to as many as 15 mn. U.S. citizens after the adoption of the Affordable Care Act (Sanger-Katz, 2015).

The evidence presented in Tables 1 to 3 is based on hypothetical rather than actual choices. Evidence from market data in the U.S. corroborates these experimental results. Studies have shown that the large degree of heterogeneity among U.S. consumers would lead to large gains in economic welfare from expanding choice (Bundorf, Levine and Mahoney 2012; Daffney, Ho and Varela 2013). Indeed, accommodating heterogeneity was one of the motivating concepts for the design of the U.S. ACA.

4 Critical interpretation of theoretical research in the light of preference heterogeneity

As shown in a companion paper (Frech and Zweifel, 2014), market socialism is well and alive in a number of sectors of western countries, among them, education and health. After the writings of Lange (1936-1938), Lerner (1934, 1936), and more recently Leeman (1977) and Roemer (1994) on market socialism, one would expect their emphasis on marginal cost pricing to be reflected in these sectors. Focusing on

health insurance, one finds this not to be the case. There, marginal cost pricing would require contributions to be graded according to the (present value of) an individual's expected future HCE, the crucial determinant of the cost associated with enrolling an additional person. Yet community rating forestalls risk-rated contributions.

However, with the exception of Roemer (1944) who discusses the problem of governance given the separation of ownership and management, the writers cited do not address the asymmetries of information that characterize health insurance. As stated above, moral hazard and adverse selection affects the relationship between the consumers and the insurer and also affects the relationship between service providers. Since the emphasis is on optimal payment of providers in the presence of heterogeneity of consumer preferences, the seminal contributions by Chakley and Malcomson (1998a, 1998b) as expounded in Zweifel et al. (2009; ch. 10.3) are reviewed here.

There is a financial sponsor (thought to be the government in Chalkley and Malcomson) who is an expected welfare maximizer. This assumption of a benevolent sponsor is unrealistic for a government agency in view of the law and economics literature (Rose-Ackerman, 1978; Shleifer and Vishny, 1993; 1994). In the case of a health insurer facing competition, it is arguably more realistic because under sufficient pressure of competition, health insurers cannot deviate much from balancing expected patient benefit from treatment against expected HCE (which determines the premium). Asymmetry of information is taken into account in that quality and treatment outcomes are considered non-verifiable. Payment P is assumed to be of the form,

$$P = F + \gamma C, \tag{1}$$

with F = a fixed component ensuring participation by the service provider and γ = degree of insurer's reimbursement of treatment cost C incurred by the provider, $0 < \gamma \leq 1$. The provider's participation constraint reads

$$EU = F - (1 - \gamma)EC(q, e) - V(q, e) = \bar{u}, \quad (2)$$

where EU = expected utility, q = quality, e = cost-reducing provider effort, EC = expected treatment cost per case, with $EC_q := \partial EC / \partial q > 0$ and $EC_e := \partial EC / \partial e < 0$, V = disutility of service provider, with $V_q < 0$ and $V_e > 0$, and \bar{u} = value of an outside option. When the first-order conditions $EU_q = 0$ and $EU_e = 0$ are incorporated to ensure incentive compatibility, the first-order condition pertaining to the optimization problem of the benevolent sponsor becomes

$$\frac{dEW}{d\gamma} = B_q(q, e) \frac{dq}{d\gamma} - \gamma EC_q(q, e) \frac{dq}{d\gamma} - V_q(q, e) \frac{dq}{d\gamma} - V_e(q, e) \frac{de}{d\gamma} = 0, \quad (3)$$

where $EW = B - \gamma EC - F$ denotes expected welfare given by patient benefit B net of provider payment according to eq. (1), B_q denotes the relationship between patient benefit and quality, $dq/d\gamma$, the provider's response in terms of quality to an increase in his or her share in cost, and $de/d\gamma$, his or her response in terms of cost-reducing effort. The terms $dq/d\gamma$ and $de/d\gamma$ are derived from comparative-static analysis of the provider's objective function (2). Throughout, $d^2EW/d\gamma^2 < 0$ is assumed.

Condition (3) can be solved for 'the' optimal value of γ . For the following reasons, however, there will be an entire set of optimal γ^* values as soon as the insured are heterogeneous. It should be noted from the outset that the terms $dq/d\gamma$ and $de/d\gamma$ are far from uniform, depending on physician characteristics such as degree of risk aversion. However, focus is on patient heterogeneity to drive home the fact that optimal payment cannot be determined without taking it into account.

1. $B_q(q, e)$, the relationship between quality and patient benefit, is likely to depend on patient type through plan selection. For instance, consider a Dutch or German who

feels subjectively ill when called upon to consider a MC-type plan. According to Table 3, he or she is likely to stick with the current plan (which entails gatekeeping but no additional restrictions in the case of the Netherlands and conventional fee-for service in the case of Germany, respectively). Such a person tends to deem an increase in quality to be especially beneficial (at a given value of q), resulting in a high value of B_q . Moreover, he or she may resent cost-reducing efforts on the part of the physician because of a belief that this diminishes the contribution of quality to patient benefit, i.e. $B_{qe} < 0$. With cost-reducing effort e at a comparatively low value in the current plan, $B_q > 0$ is augmented even more. If $dq/d\gamma > 0$ (quality enhanced by increased cost bearing by insurer, a likely effect), condition (3) states that the insurer should set γ^* at a comparatively high value in the interest of such a patient, since at a given value of $\gamma < \gamma^*$, the *EW* function of the sponsor has a more positive slope if $B_q > 0$ is large than if it is small.

2. $EC_q(q, e)$, the way expected cost per case varies with quality, is influenced by moral hazard effects on the part of the insured. If facing zero copayment as an extreme, they tend to opt for the highest quality of treatment available regardless of cost. This not only increases EC , but quite likely EC_q (at a given cost-reducing effort level e) as well because the initial level of costly quality is high already. In the case of patients facing positive copayment, their degree of risk aversion influences EC as well as EC_q . Finally, the differences in WTP values between (and within) age groups regarding the attributes ‘Physician list based on cost criteria only’ and ‘Physician list based on quality criteria’ (see Table 2) suggest that consumers generally differ in their trade-off between cost and quality of treatment, leading once again to differences in EC_q and hence γ^* .

3. $V_q(q, e)$, the degree to which the physician’s disutility of cost-reducing effort decreases with quality, is likely to reflect patient heterogeneity, too. Consider a patient

characterized by strong status quo preference being offered a new therapy for a chronic condition, such as a retired person in the Netherlands or Germany (see Table 3). This exposes the treating physician to a measure of skepticism that distracts from his or her benefit of being able to offer a higher-quality alternative, resulting in a low value of $|V_q|$. In combination with $dq/d\gamma > 0$ (see item 1 above), condition (3) calls for a low value of γ^* and hence a high provider cost share $(1-\gamma^*)$ in the interest of this type of patient. The reason is that the insurer cannot count on the provider's ethical interest in quality to rein in the moral hazard effect created by increased reimbursement of treatment cost.

By way of contrast, especially middle-aged patients appear to be, on average, strong believers in new medical technology (see the attribute 'Access to medical innovation delayed 2 years' in Table 2) who are presumably grateful to a physician proposing an innovative, quality-enhancing treatment alternative. This suggests a high value of $|V_q|$ (and possibly $B_q > 0$); with $dq/d\gamma > 0$, condition (3) calls for a high value of γ^* and hence a low provider cost share $(1-\gamma^*)$ in the interest of middle-aged insured.

4. $V_e(q, e)$, the degree to which effort directed at cost reduction increases physician disutility, also depends on patient characteristics leading to selection. For instance, according to Table 1 the attribute, 'Physician list based on cost and quality criteria' has a WTP value of -42 (14 percent of average premium, respectively) and a standard error of 7.8. Since the 20 percent premium reduction that can be legally offered for a MC-type plan corresponds to a WTP value of -29.4, which is 1.62 s.e. away from the mean value, the predicted share of Swiss consumers opting for it amounts to approximately 45 percent. Therefore, a physician having both conventionally insured and MC patients is confronted with a substantial minority of patients who expect him or her to exert effort to reduce cost as long as it does not clearly compromise quality. When treating a patient of this type, the physician experiences comparatively little

disutility when trying to achieve lower cost, implying a low value of $V_e > 0$. Combined with $de/d\gamma < 0$ (less cost-reducing effort in response to higher reimbursement of cost) condition (3) is satisfied at a low value of γ^* and hence a high provider cost share $(1-\gamma^*)$ in the case of a MC patient.

Also, a common way to achieve cost savings is to prescribe generics. According to Table 2, this is weakly resisted by the youngest age group but tends to be acceptable to the Swiss aged 65+. It therefore takes comparably little effort to convince a patient in the top age group of the equivalence between the branded and a generic drug. Given the standard assumption $V_{ee} > 0$, i.e. increasing marginal disutility of cost-reducing effort, $V_e > 0$ has a smaller value than when the patient belongs to the 25-35 age group. In view of $de/d\gamma < 0$, condition (3) calls for a high value of γ^* and hence little provider cost sharing $(1-\gamma^*)$ when it comes to prescription of generics to elderly Swiss patients.

Evidently, each one of the four components entering condition (3) gives rise to a set of different values of $(1-\gamma^*)$, the optimal degree of cost sharing by healthcare providers. This may suffice to show that optimal provider payment cannot be determined independently of insured and patient behavior, which is influenced by plan characteristics such as the rate of copayment. Since the insured are heterogeneous, both plan characteristics such as copayment and provider payment need to reflect this heterogeneity.

However, this is not possible given community rating (and more generally, premium regulation). Since downsides such as deductibles and copayment and the MC-type restrictions listed in Table 1 must not lead to differences in contributions, they cannot be compensated by the health insurer by premium reductions. Therefore, plan selection by consumers is guided predominantly by individual preferences, in particular their degree of risk aversion. These differences should optimally be reflected in the way providers are paid. However, they in turn attract differing provider types – which

the regulator is likely to view as a tool for cream skimming, justifying a refinement of RA under the impression of e.g. Glazer and McGuire (2000). Evidently, community rating causes a double welfare loss in the presence of consumer heterogeneity. Not only does it prevent health insurers from tailoring plans to individual characteristics but also structuring provider payment in the interest of their clientele.

5 Summary and conclusion

The core message of this paper is that premium regulation (in particular, community rating) in health insurance causes inefficiencies not only in the contractual relationship between heterogeneous consumers and insurers but also in that between insurers and service providers. In particular, it blocks the development of optimal provider payment, a fact that has been neglected in the existing literature. The point of departure is the observation that market socialism is prevalent in the healthcare sector with its many nonprofit and government institutions. This suggests using the prescriptions of market socialism for guidance in the pricing of goods and services. Applied to health insurance, the marginal cost pricing rule would call for contributions reflecting an extra enrollee's future expected healthcare expenditure, i.e. risk rating of contributions. Community rating disallows this, with all the consequences of deviations from the marginal cost pricing rule predicted by microeconomic theory. Closer inspection and practical experience show that risk adjustment fails to reinstate this rule.

This failure is exacerbated by consumer preference heterogeneity with regard to health insurance, for which there is substantial experimental evidence. Community rating (and premium regulation more generally) prevents health insurers from tailoring their plans to the characteristics of their clientele. Since heterogeneity implies that compensation required for accepting deductibles, rates of copayment, and MC-type restrictions differs between individuals, premiums need to be reduced to differing de-

grees. This is not compatible with community rating, causing consumers to select plans exclusively according to their individual preferences, in particular risk aversion. In addition, however, preference heterogeneity requires that the insurer structure provider payment accordingly. Taking the seminal contribution by Chalkley and Malcolmson (1998a, 1998b) as the example, it turns out that each of the four terms determining the optimal amount of provider cost sharing varies systematically with the characteristics of insured and patients. However, an insurer who structures provider payment according to the preferences of its enrollees is suspected of engaging in risk selection because this attracts certain types of providers who in turn attract certain types of patients (as argued above). If these links should result in a younger insured population, the insurer incurs a financial sanction due to risk adjustment which invariably uses age as an adjuster. Therefore, community rating causes a double welfare loss: First, consumers do not get what they want; in particular, they cannot be compensated for plan characteristics designed to limit moral hazard effects. Second, the insurer has little incentive to structure provider payment optimally reflecting the heterogeneity of its clientele.

The lesson for policy is straightforward [see also Zweifel and Pauly (2007), Bundorf et al. (2010), and Bhattacharya et al. (2013) as well as Frech and Zweifel (2014)]. Health insurers both private and social should be encouraged to risk-rate contributions, resulting in marginal cost pricing of insurance coverage. In a competitive market equilibrium, profit margins do not differ between high and low risks because there cannot be cross-subsidization under the pressure of competition. This means that insurers have no reason to prefer one type over the other, annihilating their incentive to invest in risk selection or to distort their offerings to select favorable risks. In a multi-period setting, they can overcome problems of informational asymmetry by using loss experience to determine risk types and withdrawing loss-making contracts. In addition, they can implement experience rating of contributions (also known as bonus options), which have been found to rein in moral hazard effects (Zweifel, 1987). Given risk-rating of contributions, concerns about risk selection do not militate any more against structuring provider payment in the interest of different consumer groups

characterized by differences in preferences. If this results in some healthcare providers bringing in groups of insured with low healthcare expenditure relative to premium paid, they are competed away by health insurers offering them a lower premium. This process comes to an end when profit margins are equalized across types of consumers again.

The standard criticism of such a proposal is that some high risks cannot afford the risk-rated premium. While this is true, it does not justify premium regulation from an economic point of view. As stated in the Second Theorem of Welfare Economics, the efficient solution is to encourage competition to drive prices towards marginal cost while modifying the initial income distribution by taxes and transfers. Therefore, those high risks who cannot afford the premium should receive a tax-financed subsidy, to be targeted not at high risks in general but at high risks of modest means. An annual contribution of \$ 20,000 (say) is affordable to a consumer who earns an income of \$ 200,000 per year. Whether it is deemed affordable for someone earning \$ 100,000 is a political decision, to be made by parliament in a representative democracy and voters in a direct democracy. Switzerland provides an illustration of the latter case. In December 1995, the new law on health insurance survived a popular referendum -- not least (according to surveys) because it provided for a subsidy for those whose premium would exceed between eight and ten percent of taxable income (the precise choice of threshold was left to the cantons).

There are downsides to this efficient solution. First, the premium subsidy is transparent. Its cost can be seen in the government's budget, while the cost of premium regulation is hidden. Economists view this transparency as a benefit. For example, in a classic work, Pauly et. al. (1992) suggest a similar policy, arguing that

(The plan) has a very strong advantage in encouraging rational democratic political choice—an advantage to voters, though not necessarily to politicians...The trade-offs between beneficiaries

and taxpayers will be obvious and subject to discussion and will likely be decided in open public debate. (Pauly et. al. 1992, p. 21)

On the other hand, as noted in the citation, transparency may be disliked by politicians who prefer to confer visible benefits to their constituencies while burdening the remainder of society in way that are not easily recognized. Moreover, premium subsidies may

lead to problems of implementation. Political attention may cause the subsidy to become poorly targeted. In some cantons of Switzerland, up to 50 percent of households receive a premium subsidy, apparently because local governments sought to ensure their re-election by lowering the threshold below eight percent of taxable income or by narrowly defining 'taxable income'.

Also, in some countries several social benefits (housing, college support, family allowances) are means-tested. Means-tested benefits create disincentives for labor. Tying a subsidy for health insurance to the same income threshold increases the disincentives for labor supply.

However, the efficiency advantages of risk-rated contributions to health insurance complemented by targeted subsidies seem to be important enough to merit consideration. After all, health insurance improves access to medical care – arguably one of the most personal services there are. Therefore, a policy imposing 'one size fits all' in health insurance must burden the economy with substantial welfare losses. Avoiding them is worth economic argument based on continuing theoretical and empirical research.

References

- Beck, K., Zweifel, P. and Trottmann, M. (2010), Risk adjustment in health insurance and its long-term effectiveness, *Journal of Health Economics* 29(4), 489-498.
- Bhattacharya, J., Hyde, T., and Tu, P. (2013), *Health Economics*. London: Palgrave Macmillan.
- Bhattacharya, J. et. al. (2013), *Best of Both Worlds: Uniting Universal Coverage and Personal Choice in Health Care*, Washington, D.C.: American Enterprise Institute Press.
- Bundorf, M. K., Herring, Bradley, H. and Pauly, M.V. (2010), Health risk, income, and employment-based health insurance. *Forum for Health Economics & Policy* 13(2), 1-35.
- _____, Levine, J., and Neale, M. (2012), Pricing and welfare in health plan choice. *American Economic Review* 102(7), 3214-3248.
- Chalkley, M. and Malcomson, J. (1998a), Contracting for health services when patient demand does not reflect quality. *Journal of Health Economics* 17, 1-19.
- Chalkley, M. and Malcomson, J. (1998b), Contracting for health services with unmonitored quality. *Economic Journal*, 1093-1110.
- Clemens, Jeffrey (2015), Regulatory Redistribution in the Market for Health Insurance, *American Economic Journal: Applied Economics* 7 (2), 109-134.

- Dafny, Leemore, Kate Ho and Mauricio Varela (2013), Let Them Have Choice: Gains from Shifting Away, from Employer-Sponsored Health Insurance and Toward an Individual Exchange, *American Economic Journal: Economic Policy* 5 (1), 32-58.
- Eisenhardt, K. (1989), Agency theory: an assessment and review. *Academy of Management Review* 14(1), 57–74.
- Feldstein, M. (2013), *Remarks at the American Economic Association Meetings Honoring Amy Finkelstein*, Philadelphia, PA.
- Frech, H.E., III and Smith, Michael P. (2015), Anatomy of a Slow-Motion Health Insurance Death Spiral, *North American Actuarial Journal*, 19 (1), 60-72.
- Frech, H.E. III and Zweifel, P. (2014), *Market Socialism and Community Rating in the Affordable Care Act*. Santa Barbara CA: UCSB Discussion Paper.
- Glazer, J. and McGuire, T. (2000), Optimal risk adjustment in markets with adverse selection: an application to managed care. *American Economic Review* 90(4), 1055-1071.
- Halek, M. and Eisenhauer, J.G. (2001), [Demography of risk aversion](#). *Journal of Risk and Insurance* 68(1), 1 -24.
- Jehle, G.A. and Reny, P.J. (2001), *Advanced Microeconomic Theory*, 2nd. ed. Boston: Addison-Wesley.
- Lancaster, K.J. (1966), A new approach to consumer theory. *Journal of Political Economy* 74 (2), 132-157.
- Lange, O. (1936), On the economic theory of socialism. *Review of Economic Studies* 4 (1-2) (Oct. 1936, Feb. 1937), 53-71 and 123-142, reprinted with small changes in B. E. Lippincott (Ed.), *On the Economic Theory of Socialism*. University of Minnesota Press (1938), 55-143.
- Leeman, W. A. (1977), *Centralized and Decentralized Economic Systems*. Chicago: Rand McNally.

- Lerner, A. P. (1934), Economic theory and socialist economy. *Review of Economic Studies* 2(1), 51-61.
- Lerner, A. P. (1936), A note on socialist economics. *Review of Economic Studies* 4(1), 72-76.
- Leukert-Becker, K. and Zweifel, P. (2014), Preferences for health insurance in Germany and the Netherlands – a tale of two countries. *Health Economics Review* 4(1).
- MacNeil Vroomen, J. and Zweifel, P. (2011), Preferences for health insurance and health status: does it matter whether you are Dutch or German?. *European Journal of Health Economics* 12(1), 87-95.
- Sanger-Katz, M. (2159), New York Times (2015), High Rate of Shopping and Switching in Obamacare Plans Is a Good Sign, 26 Feb.
- Pauly, M.V. (1970), The welfare economics of community rating. *Journal of Risk and Insurance* 37(3), 407-418.
- Pauly, M.V., et. al. (1992), *Responsible National Health Insurance*. Washington, D.C.: American Enterprise Institute Press
- Roemer, J. E. (1994), *A Future for Socialism*. Cambridge, Mass.: Harvard University Press.
- Rose-Ackerman, S. (1978), *Corruption: A Study in Political Economy*. New York: Academic Press.
- Schoder, J., Sennhauser, M., and Zweifel, P. (2010), Fine-tuning of health insurance regulation – unhealthy consequences for an individual insurer. *International Journal of the Economics of Business* 17(3), 313-327.

- Shleifer, A. and Vishny, R. W. (1993), Corruption. *Quarterly Journal of Economics*, 108(3), 599-618.
- Shleifer, A. and Vishny, R.W. (1994), The Politics of Market Socialism, *Journal of Economic Perspectives*, 8(2), 165-176/
- Tinbergen, J. (1954), *On the Theory of Economic Policy*. Amsterdam: North-Holland.
- Zweifel, P. (1987), Bonus systems in health insurance: a microeconomic analysis. *Health Policy* 7, 273-288.
- Zweifel, P. (2013a), The present state of health economics: a critique and an agenda for the future (Editorial), *European Journal of Health Economics* 14(4), 569-571.
- Zweifel, P. (2013b), Managed Care: prescription for failure? Lessons from Switzerland, *Cesifo DICE Report (Journal for Institutional Comparisons)* 11(1), 27-31.
- Zweifel, P. and Manning, W.G. (2000), Moral hazard and consumer incentives in health care. In: A.J. Culyer and J.P. Newhouse (Eds.), *Handbook of Health Economics*, Vol. 1. Boston: North Holland, 409-459.
- Zweifel, P. and Breuer, M. (2006), The case for risk-based premiums in public health insurance, *Health Economics Policy and Law*, 1(2), 171-188.
- Zweifel, P., Telser, H., and Vaterlaus, S. (2006), Consumer resistance against regulation: the case of health care. *Journal of Regulatory Economics* 29 (3), 319-332.
- Zweifel, P. and Pauly, M.V. (2007), Market outcomes, regulation, and policy recommendations. In: A. S. Preker et al. (eds.), *Private Voluntary Health Insurance in Development – Friend or Foe?* Washington: The World Bank, 115-145.
- Zweifel, P., Breyer, F. and Kifmann, M. (2009), *Health Economics*, 2nd ed. New York: Springer.