Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

ISSN 2282-6483

Prevention in Health Insurance: a Welfare Analysis of Participating Policies

Francesca Barigozzi
Renaud Bourlès
Dominique Henriet
Giuseppe Pignataro

Quaderni - Working Paper DSE N°765

Prevention in Health Insurance: a Welfare Analysis of Participating Policies

Francesca Barigozzi[a*] Renaud Bourlès[b†] Dominique Henriet[b‡]
Giuseppe Pignataro[a§]

[a] *Department of Economics, University of Bologna, P.zza Scaravilli 2, 40126 Bologna, Italy*

[b] *Ecole Centrale Marseille (Aix-Marseille School of Economics), CNRS & EHESS*

38 rue Frédéric Joliot-Curie, 13451 Marseille Cedex 20, France

July 2014

Abstract

Preventive care should be subsidized in traditional insurance contracts since policyholders ignore the benefit of their prevention choice on the insurance premium (Ellis and Manning, 2007 JHE). We study participating policies as risk-sharing agreements among policyholders who decide how much to invest in secondary prevention. We explore under which conditions these policies allow partial or even full internalization of prevention benefits in an environment with repeated interactions between policy holders. Welfare generated by the risk-sharing agreement is increasing with the size of the pool, but at the same time the pool size must not be too large for cooperation to sustain the internalization benefits.

Key words: secondary prevention, positive externality on the insurance premium, long run enrollment, cooperation among policyholders.

1 Introduction

All developed countries are deeply engaged in improving prevention strategies and prevention policies to protect, promote and maintain health and to prevent disease, disability and premature death. For example, according to the US Center for Disease Control and Prevention: "Called for by the Affordable Care Act, the National Prevention Strategy includes actions that public and private partners can take to help Americans stay healthy and fit. It helps move the nation away from a health care system focused on sickness and disease to one focused on wellness and prevention."[1] The effectiveness of disease prevention programs depends, among other things, on incentives provided by the insurance system and on the resulting consumer choices.

*E-mail address: francesca.barigozzi@unibo.it

†E-mail address: renaud.bourles@centrale-marseille.fr

‡E-mail address: dominique.henriet@centrale-marseille.fr

§E-mail address: giuseppe.pignataro@unibo.it

[1] US Center for Disease Control and Prevention, http://www.cdc.gov, consulted in March 2013.

Some papers have studied optimal health insurance contracts when policyholders choose consumption of both preventive care and treatment. Ellis and Manning (2007) investigate optimal coverage for primary prevention, showing that a positive copayment for preventive care is desirable since traditional cost-sharing contracts ignore the beneficial impact of preventive care on the insurance premium. In other words, since the insurance premium is fixed and is not affected by the policyholders' prevention choices, without a positive coverage consumers would under-invest in prevention. In a similar way, when secondary prevention and treatment are substitute inputs in the health recovery function, Barigozzi (2004) shows that a positive copayment for secondary prevention is optimal.

The aim of this paper is to analyze alternative contracts encouraging prevention: participating policies. As we will show, participating policies can lead to full internalization of the positive effect that prevention has on the insurance premium. This result is basically driven by two main facts: (i) participating policies induce more responsible (or less myopic) behaviors since policyholders bear the aggregate health risk; (ii) firms offering participating policies generally promote long run relationships with their policyholders, and long run enrollment proves to be effective in increasing preventive efforts.

When purchasing participating policies[2], consumers jointly hold the residual claims of the pool. In practice, policyholders become the members of a risk-sharing agreement and contribute whatever amount is needed yearly to meet the health expenditures covered by the policy (Doherty and Dionne 1993, Picard 2009). Risk-sharing is usually in the form of an initial contribution followed by later "calls", if required, to maintain the common fund. More importantly, the premium of a participating policy is random (since it depends on how many and how important negative health shocks are finally realized) and, as a consequence, policyholders in risk arrangements always face aggregate risk.

Since participating policies have been, and currently are, mainly offered by mutual insurance firms, many studies in the literature use the terms 'mutual' and 'participating policies' interchangeably (see Smith and Stutzer, 1995). We investigate participating policies, no matter what type of firm is offering them, but we will frequently refer to empirical evidence on mutuals insurers and to mutuals' corporate culture and mission to support our arguments and results.

Our paper offers a welfare analysis of participating policies in health insurance when prevention matters. In particular, we investigate *secondary* prevention which refers to all measures allowing early detection of disease such as diagnostic screening, medical examinations and checkups.[3] In the past decade, the policy debate on secondary prevention has become particularly active. For example in May 2005 the 58th World Health Assembly of the World Health Organization approved a resolution calling on Member States to intensify compliance with cancer prevention by encouraging screening tests such as mammograms, pap smears and colonoscopy. The design of incentive compatible insurance policies is obviously a crucial instrument that policy-maker can use to increase compliance with preventive care.

We study participating policies as risk-sharing agreements among identical policyholders who decide how much to invest in secondary prevention. In particular, secondary prevention is modeled as a costly action that, when illness occurs, reduces health care expenditures necessary to recover.

[2] Participating policies are available for all types of insurance coverage: health, medical malpractice, life, car, fire, agriculture, maritime insurance, etc.

[3] According to the health economics literature, primary prevention reduces the *probability* of illness, whereas secondary prevention reduces *vulnerability* to illness. In other words, primary prevention, such as physical exercise and diets, concerns the avoidance of undesirable outcomes. Secondary prevention measures instead, by detecting an illness when it is still asymptomatic, reduce its incidence and facilitate patient recovery (see Kenkel 2000).

In the model, policyholders can choose the level of prevention either non-cooperatively or cooperatively. Non-cooperative choice allows for *partial* internalization of the effect of prevention on the insurance premium, whereas cooperation allows for *full* internalization and thus implies an efficiency gain. We find that partial coverage is optimal under the non-cooperative strategy while full coverage is optimal whenever policyholders cooperate. Moreover, the cooperative outcome always dominates the non-cooperative equilibrium in terms of policyholders' welfare and replicates the first-best allocation as pool size rises to infinity.

However, participating policies do not always deliver cooperation. We identify conditions under which cooperation can be sustained as an equilibrium within a repeated interaction game. Since policyholders' incentives to free ride are always higher in a large pool, the equilibrium with cooperation, if it exists, is enforceable only when the pool size is not too high. When the pool size is so large that the cooperative equilibrium is not enforceable, then the non-cooperative equilibrium necessarily arises. This depicts a schedule for policyholders' welfare in equilibrium that is piecewise increasing in the pool size, but discontinuous (with a downward jump when cooperation turns out to be no longer enforceable, see Figure 3). We conclude that, in risk sharing arrangements, policyholders may fully internalize the benefit of their prevention choice on the insurance premium. However a trade-off arises with respect to the number of policyholders purchasing the policy. On the one hand, the benefit from risk-sharing is increasing in the size of the pool; on the other hand efficiency (that is full internalization of the impact of prevention) is only compatible with a pool size that is not too large.

Interestingly, some evidence confirms the relatively small dimension of mutual companies, especially in Europe. For example, according to the International Cooperative and Mutual Federation (ICMIF)[4], the total market share of the 2,900 mutual firms active in 75 countries at the end of 2010 was just 26%. More detailed evidence can be found in France and Italy. Caire (2009) reports that, in 2007 in France, the 808 existing mutuals accounted for 58% of the health insurance market, the 9 active stock (i.e. standard) insurers owned 23% of the market while the remaining 19% was the share of "institutions de prévoyance" (non-profit organizations that offer collective insurance contracts for firms).[5] Likewise in Italy, health mutual firms are definitely characterized by small size: about 1,500 "Società di mutuo soccorso" (mutual benefit societies) were active in 2010, they had less than one million members and represented around 12% of the Italian market in complementary health insurance (see Lippi Bruni *et al.* 2012).

Our results are also in line with the common view that participating policies in small pools allow better coordination of policyholders' behaviors and with both empirical and theoretical literature on organizational form in insurance and risk-sharing (see for example Ligon and Thistle, 2005; Fafchamps and Lund, 2003; Genicot and Ray, 2003 or Bramoullé and Kranton, 2007).

Our model shows that cooperation can be obtained when policyholders play a repeated game with the other members of the risk-sharing agreement. Thus, cooperation is the outcome of a long run enrollment. Importantly, while no incentives to a long run interaction (that makes prevention profitable for the insurer) are offered to policyholders in standard health policies, firms offering participating policies seem to encourage long run enrollment and cooperative behaviors in different ways.

In this perspective, Goeffard (2000) shows that in France premiums of mutual insurers increase

[4]See the Global Mutual Market Share report 2010, available at http://www.icmif.org/mms2010.

[5]Similarly, Kerleau (2009) shows that the market for participating policies is characterized by low concentration in France, as the 5 biggest mutuals in 2005 represented only 20% of the market share and the 30 biggest ones only 44%.

less with age than those of stock insurers. Broek *et al.* (2012) compare the different national legal regimes for mutuals in Europe. On the one hand mutual companies are built on solidarity and mutual support principles irrespectively of the risk they cover; on the other hand, mutuals providing health coverage seem naturally committed to prevention, given their interest in long-term contracts. For instance, the French "Code de la mutualité" regulates all aspects of the organization of mutuals, carrying out solidarity, mutual aid-based work and emphasizing a culture of healthy life-style and preventive behaviors.[6]

More generally, firms offering participating policies seem to provide effective incentives to their policyholders and are traditionally associated with lower costs.[7] Indeed, the mutual/cooperative sector performed relatively better during the recession: at a world level the mutual/cooperative market share increased by 2.7% from 2007 and it increased by 5.2% in Europe. In this respect, it is argued that the better-than-market performance of mutuals may have been boosted by higher levels of customer trust and customer satisfaction associated with the participating form of the policy, which may easily translate into higher levels of cooperative behaviors.

Our paper is also related to the insurance literature studying the efficiency of risk-sharing arrangements. The latter have been diffusely analyzed, for example, in settings with asymmetric information (see, among others, Mayers and Smith 1986 and Smith and Stutzer 1990). In this literature, the contribution most closely related to our paper is Lee and Ligon (2001). They study optimal risk-sharing contracts when policyholders use a non-cooperative strategy in the choice of a self-protection measure, while we analyze self-insurance (secondary prevention) focusing on cooperative behaviors as possible alternatives to non-cooperative ones.

The structure of the article is described below. Section 2 illustrates the model set-up. Then two benchmark cases are described: the first-best in Subsections 2.1 and, the second-best policy with fixed premium in Subsection 2.2. Section 3 introduces participating policies. Sections 4 and 5 study the non-cooperative equilibrium and the cooperative outcome in the risk-sharing agreement, respectively. The relationship between pool size and efficiency is discussed in Section 6. Section 7 shows how cooperation can be enforced. Section 8 discusses some policy issues. Concluding remarks follow in the last section.

2 The model

The economy is composed of n identical individuals with utility from money represented by a strictly increasing and concave von Neumann-Morgenstern utility function $U(w)$ which is differentiable at least twice. Individuals have initial wealth w and face the probability p of a monetary loss of size $L(a)$ with independently and identically distributed risks.

The loss $L(a)$ is a function of individuals' action a such that $L'(a) < 0$. From Ehrlich and Becker (1972), a consumer's action decreasing the amount of a possible loss is a self-insurance measure. By interpreting $L(a)$ as the monetary equivalent of a negative health shock, the action a refers to secondary prevention or early detection of disease. Secondary prevention is exerted before the risk is realized. As specified in the introduction, medical examinations, check-ups and diagnostic screening help to reduce the severity of illness, and as such are to be considered secondary prevention measures. In particular, secondary prevention decreases the amount of health care necessary to recover from the negative health shock.

[6] See also our discussion in Section 8.

[7] See Mayers and Smith (1986), Cummins and Zi (1998) and Cummins et al. (1999), (2004) for studies documenting the better performance of mutuals in the life insurance market.

The disutility from secondary prevention $C(a)$, with $C'(a) \geq 0$ and $C''(a) > 0$, is due to time-consuming, invasive and sometimes unpleasant diagnostic procedures (e.g. colonoscopy.), and is assumed to be additively separable from the utility derived from money.[8] To assure an internal solution for prevention choice we also impose $C'(0) = 0$.

A policy (partially) reimburses the losses suffered by policyholders in exchange for an insurance premium. Although for the sake of tractability the function $L(a)$ is deterministic,[9] the insurer cannot perfectly control secondary prevention (so that moral hazard will be an issue) because it offers cost-sharing or linear contracts. More precisely, policyholders obtain a partial reimbursement of their health care expenditure according to the fraction q, $0 \leq q \leq 1$, so that the contract pays $qL(a)$ in case of illness. Cost-sharing contracts are also the prevailing type of contract in health insurance.

Prevention reduces the amount of treatment necessary to recover so that health expenditure (and thus reimbursement paid by the insurer) decreases. We will show that, with participating policies, policyholders might fully internalize the benefit of their prevention choice on the premium. Before doing so, in the next two sub-sections we analyze two benchmark cases, useful for further comparisons: the first-best contract and the second-best one. The latter, characterized by a fixed premium, represents a non-participating policy contract in outline.

2.1 The first-best contract

We consider here optimal coverage for a representative consumer in a perfect information environment. The first-best contract solves the following program:

$$\max_{a,q} EU = pU\left(w - pqL(a) - L(a) + qL(a)\right) + (1-p)\,U\left(w - pqL(a)\right) - C(a) \tag{1}$$

where policyholders receive $qL(a)$ in the event of illness and the fair premium is $pqL(a)$.

Let's denote net consumption in the two states of nature as W_L when loss occurs and W_0 with no loss, respectively. The optimal value of the cost-sharing parameter q is full coverage $\left(q^{FB} = 1\right)$, so that net consumption in the two possible states of the world is the same: $W_L = W_0 = W = w - pL(a)$. Consequently, the optimal choice of prevention is:

$$a^{FB} : U'(w - pL(a))\,(-pL'(a)) = C'(a). \tag{2}$$

The left-hand side of (2) shows the marginal benefit while the right-hand side shows the marginal cost of prevention. Note that, in the first-best contract, policyholders perfectly internalize the beneficial effect of prevention on the premium (see the term $U'(w - pL(a))$). In particular, they take into account that a higher level of self-insurance, by decreasing the premium, has a positive impact on marginal utility in both the possible states of nature. Marginal benefit is increasing in p and in $-L'(a)$, i.e. the efficiency of the prevention technology. Policyholders' welfare is maximized and corresponds to:

$$EU^{FB} = U\left(w - pqL(a^{FB})\right) - C(a^{FB}) \tag{3}$$

[8] Secondary prevention could also imply a monetary cost to be incorporated in the loss function $L(a)$. This would not affect our results, provided that the property $L'(a) < 0$ is maintained or that the advantage of prevention more than compensates its cost.

[9] In the case of participating policies, the premium depends on the number of policyholders suffering the loss, which is unknown ex-ante, so that expected utility must be calculated with a complex binomial function (see expressions 8 and 11 below). Hence, assuming a stochastic relationship between a and L would dramatically complicate the analysis without adding relevant insights.

2.2 The second-best contract with fixed premium

As mentioned above, we interpret the second-best contract as a non-participating policy with fixed premium. The timing of actions is the following: first, the insurer proposes the contract; second, the consumers accept the contract and choose the prevention level; finally the risk is realized. Policyholders receive $qL(a)$ in the event of illness and the premium is $P = pqL(a)$. A fair premium is coherent both with the case of a benevolent monopolistic health insurer (i.e. a public/social insurance) and with the case of a large number of insurers in a competitive market. The representative policyholder's expected utility is:

$$EU^{SB} = pU\left[w - L(a) - P + qL(a)\right] + (1-p)\,U\left(w - P\right) - C(a) \tag{4}$$

The optimal prevention level is calculated *given* the contract (P, q). The policyholder's optimal choice thus verifies:

$$a^{*SB}(q) : -(1-q)L'(a)pU'(W_L) = C'(a) \tag{5}$$

Obviously, if $q = 1$ then no prevention is the outcome so that the insurer will never offer full insurance and $W_L \neq W_0$. By comparing (2) and (5) we observe that in the latter FOC policyholders do not definitively internalize the positive impact that prevention has on the premium, as precisely emphasized by Ellis and Manning (2007). In the l.h.s. of (5) only the beneficial effect of prevention on the potential illness (when net consumption is W_L) is taken into account. We will show that the FOC is different in the case of a participating policy since the policyholders (at least partially) internalize the positive impact of prevention on the premium.

To derive the optimal second-best contract we proceed backward: the insurance firm maximizes the policyholder's utility (4) subject to the resources constraint ($P = pqL(a)$) and the policyholder's incentive constraint (5). Solving this program, the optimal level of coverage q is found to be lower than 1 (partial coverage), which means that the usual trade-off between risk-sharing and incentives arises.

3 Participating policies

We are going to define participating policies in the case of cooperative and non-cooperative choice of prevention among policyholders. Then we will compare outcomes under the two different strategies both for a finite pool size and when the number of policyholders goes to infinity. Those comparisons are necessary in order to set up the analysis of cooperation among members of the risk-sharing arrangement.

Suppose now that the insurer offers a participating policy to the n identical individuals. Again q is the percentage of the loss reimbursed to each policyholder and $qL(a)$ is the indemnity received in the event of illness. Let's call K the number of policyholders that experience the negative health shock out of the n identical individuals in the pool: $K \in \{0, ..., n\}$. The peculiarity of the participating policy is that the *aggregate* indemnity reimbursed to the policyholders, and hence the required individual premium, are not fixed; both depend on the realization of K. This implies that the individual premium is *random*.

Definition 1 *The participating policy is such that the aggregate amount of indemnities to be paid to policyholders belonging to the pool ($KqL(a)$) is equally shared among the n members of the pool.*

Thus, the individual premium is: $\frac{KqL(a)}{n}$[10].

The overall amount of premiums collected by the pool exactly covers the aggregate indemnities paid to the K individuals experiencing the negative health shock. This is a standard property of participating policies: profits are always zero ex-post[11].

The timing of actions is the following:

1. the coverage level, q, is chosen by the pool.

2. Each policyholder chooses his prevention level either non-cooperatively or cooperatively.

3. The risk (and thus the number K of policyholders submitting a claim) is realized.

Note that, in the second stage, policyholders' choice of prevention can be either cooperative or non-cooperative.

3.1 Non-cooperative choice of prevention

We now investigate the case where prevention is chosen non-cooperatively. The representative policyholder only internalizes the effect *of his own prevention* on the random premium in the event he experiences the loss. In other words, the policyholder neglects the "social" benefit of prevention on the aggregate loss in both states of nature. Remember that K is the total number of policyholders experiencing the loss in the pool. Here we take into account the view of the representative policyholder and we define k as the number of individuals that experience the loss once we exclude the representative consumer from the pool. This allows us to write the expected utility of the representative policyholder. If the policyholder i experiences the illness himself, then $K = k + 1$ whereas if he does not, $K = k$. Thus, for a given realization of k:

$$\begin{aligned} EU_i^{NC}\left[a_i; n, q, a_{-i} \,|k\right] &= pU\left(w - \frac{q}{n}\left(L(a_i) + kL(a_{-i})\right) - L(a_i) + qL(a_i)\right) \\ &+ (1-p)\,U\left(w - \frac{kq}{n}L(a_{-i})\right) - C(a_i) \end{aligned} \tag{6}$$

where a_i is the prevention level exerted by i and a_{-i} is the level exerted by the other $n-1$ policyholders in the pool.

It is worth noting that policyholder i behaves as if the impact of his own prevention on the premium were different in the two states of nature. In particular, he perceives that his prevention level affects the insurance premium only when he suffers the negative health shock, such that the premium writes $\frac{q}{n}\left(L(a_i) + kL(a_{-i})\right)$ in this case, whereas he perceives that his premium does not depend on his prevention level when he is healthy, such that the premium writes $\frac{kq}{n}L(a_{-i})$ when no loss occurs.

To summarize, in the non-cooperative case, each policyholder internalizes part of the effect of his own prevention on the premium, i.e. $\frac{q}{n}L(a_i)$, and only in the event of illness.

[10] Note that such an "equal sharing rule" is not the optimal incentive-compatible rule. The optimal rule would be obtained by maximizing the expected utility of a representative policyholder under the incentive constraint, and a resource constraint that would have to be fulfilled in every state of nature (and not only in expectation). Still, the equal sharing rule defined above seems to better describe actual participating policies, in particular in the health insurance market.

[11] In the case of participating policies, if at the end of the period the aggregate indemnities to be reimbursed by the insurer are greater than the premiums collected, policyholders are asked to pay an additional premium. If, on the other hand, the aggregate indemnities to be reimbursed are lower than the premiums collected, either policyholders receive money back or the insurer uses "profits" to increase its contingency reserves and funds.

3.2 Cooperative choice of prevention

We innovate with respect to the existing literature on risk-sharing agreements by allowing policyholders to choose the prevention action cooperatively, thus making it possible to fully internalize the "social" benefit of the policyholders' prevention choice on the aggregate health expenditures.

Definition 2 *Cooperation is the situation where all policyholders agree on a common prevention level that maximizes the expected welfare of a representative policyholder.*

Therefore, we consider here the representative policyholder's problem as his choice of the cooperative prevention level when all the other policyholders cooperate as well. The representative policyholder's expected utility, given that k members out of the others $(n-1)$ experience the illness, is now:

$$EU^C\left[a;n,q\,|k\right] = pU\left(w - L(a) - \frac{(k+1)qL(a)}{n} + qL(a)\right) + (1-p)\,U\left(w - \frac{kqL(a)}{n}\right) - C(a) \quad (7)$$

where a is the prevention level chosen by all cooperating consumers in the pool. Cooperation allows policyholders to fully internalize the positive externality exerted by their prevention choice on the insurance premium. Thus, we expect a higher level of prevention under cooperation than in the case of non-cooperation.

In the following sections, we will first analyze the non-cooperative game, and then we will derive the policyholders' payoff under cooperation.

4 Non-cooperative equilibrium

When policyholders do not cooperate, the representative individual i's expected utility, given that k other members of the pool have experienced the illness, is expressed in (6) above. Considering all possible realizations of k :

$$EU_i^{NC}(a_i;n,q,a_{-i}) = \sum_{k=0}^{n-1} b(k;n-1;p)EU_i^{NC}\left[a_i;n,q,a_{-i}\,|k\right], \quad (8)$$

where $b(k;n-1;p)$ is the binomial probability of k negative health shocks with $n-1$ individuals characterized by probability of illness p.

Solving backward, in the second step the representative policyholder chooses his own prevention level. The optimal choice maximizes the expected utility taking as given the prevention level chosen by other policyholders a_{-i} and the coverage q, that is $a_i^{*NC}(n,q,a_{-i}) = \arg\max_{a_i}$ $EU_i^{NC}(a_i;n,q,a_{-i})$. In particular:

$$a_i^{*NC}(n,q,a_{-i}) : \sum_{k=0}^{n-1} b(k;n-1;p)\{pU'\left[W_L^{NC}\right](-(1-q)L'(a_i) - \frac{q}{n}L'(a_i))\} = C'(a_i) \quad (9)$$

where, as previously explained, net consumption in the case of illness is $W_L^{NC} = w - \frac{q}{n}\left(L(a_i) + kL(a_{-i})\right)$ $-L(a_i) + qL(a_i)$.

Interestingly, from (9) and contrary to (5), we see that with a participating policy, prevention would be positive even in the case of full coverage ($q = 1$), since the marginal benefit of prevention is always greater than zero. The reason is that here policyholders internalize part of the beneficial impact of prevention on the premium through the term $\frac{q}{n}L(a_i)$ appearing on the left-hand side of equation (9). However, we will see below that full coverage is never optimal when policyholders *do not* cooperate, just as in the case of the second-best non-participating policy analyzed in Subsection 2.2.

With identical agents, the equilibrium is symmetric and $a_i = a_{-i} = a$ so that:

$$a^{*NC}(n,q) : \sum_{k=0}^{n-1} b(k; n-1; p)\{pU'\left[W_L^{NC}\right](-L'(a)(1-q+\frac{q}{n}))\} = C'(a) \tag{10}$$

where $W_L^{NC} = w - \frac{(k+1)q}{n}L(a) - (1-q)L(a)$.

We can now consider the first step of the game: since the coverage is chosen collectively and is unique in the pool, q^{*NC} is the solution of the program where the expected utility of the representative policyholder is maximized with respect to coverage q under the incentive compatibility constraint (10).

Lemma 1 *When policyholders act non-cooperatively in the second stage, the participating policy is characterized by partial coverage* $\left(q^{*NC} < 1\right)$.

Proof. See the Appendix 10.1. ■

Similarly to the second-best contract considered in Subsection 2.2, when they choose prevention non-cooperatively, policyholders receive here a partial reimbursement. This result holds even though, given the participating policy and the random premium, policyholders face a higher risk than in the second-best. For this reason we expect that the optimal (partial) coverage within the participating policy will be higher than the optimal (partial) coverage in second-best.[12]

5 The cooperative outcome

Under cooperation, the representative policyholder i's expected utility, given that k individuals other than i have experienced the loss, is expressed in (7) above. Considering all possible values of k, expected utility becomes:

$$EU^C(a; n, q) = \sum_{k=0}^{n-1} b(k; n-1; p) EU^C\left[a; n, q \,|k\right]. \tag{11}$$

In the second step of the game, the optimal choice of prevention, given the risk-sharing rule and

[12] Lee and Ligon (2001) analyze the non-cooperative solution in a mutual arrangement when a self-protection measure is available to policyholders (i.e. in the case of an action that decreases the loss probability). Using "Cournot conjectures" they show that, despite the presence of ex-ante moral hazard, full coverage is optimal. Our result differs from theirs since we solve the problem using the concept of Nash equilibrium. This implies that, when choosing the reimbursement level q in the first step, policyholders anticipate the effect of such a coverage on the prevention choice of other members of the pool. Thus, the second term in expression (19) of the proof, in Appendix 10.1, is different from zero in our model, whereas the corresponding term is zero in Lee and Ligon (2001).

the coverage q, is $a^{*C}(n,q) = \arg\max_a EU^C(a;n,q)$. In particular:

$$a^{*C}(n,q):$$
$$\sum_{k=0}^{n-1} b(k;n-1;p)\Big\{p(U'\left[W_L^C\right](-L'(a)+\tfrac{(n-(k+1))}{n}qL'(a)) \tag{12}$$
$$+(1-p)U'\left[W_0^C\right]\left(-\tfrac{kqL'(a)}{n}\right)\Big\} = C'(a)$$

where net consumption can be written as: $W_L^C = w - \frac{(k+1)q}{n}L(a) - L(a) + qL(a)$ and $W_0^C = w - \frac{kq}{n}L(a)$.

We can now consider the first step of the game: under cooperation the optimal coverage q^{*C} solves the program where expected utility $EU^C(a;n,q)$ is maximized with respect to the coverage q under the incentive compatibility constraint (12).

Corollary 1 *When policyholders act cooperatively in the second stage, the participating policy implies full coverage* $\left(q^{*C}=1\right)$.

Proof. See Appendix 10.2. ■

Note that we obtained partial coverage under the non-cooperative strategy but full coverage in the case of cooperation. Intuitively, under cooperation, policyholders fully internalize the impact their action has on the premium and thus choose a sufficiently high level of prevention, even when the coverage is 1. In the non-cooperative case, on the other hand, partial internalization implies that policyholders need higher incentives, in the form of partial coverage, to choose a sufficiently high level of secondary prevention[13].

From (12), one can easily check that the optimal action is positive even with full coverage. In particular, when $q^{*C}=1$, the optimal prevention becomes:

$$a^{*C}(n,1):$$
$$\sum_{k=0}^{n-1} b(k;n-1;p)\Big\{-\left(\tfrac{1}{n}\right)L'(a)pU'\left(w-\frac{k+1}{n}L(a)\right) - E\left[U'(W^C)\right]L'(a)\tfrac{k}{n}\Big\} = C'(a) \tag{13}$$

where $E\left[U'(W^C)\right] = pU'\left(W_L^C\right) + (1-p)U'\left(W_0^C\right)$.

Concerning the implicit solutions for the policyholders' optimal prevention level, i.e. equations (5), (10) and (13), we note that the term $-E\left[U'(W^C)\right]L'(a)\frac{k}{n} > 0$ only appears in the left-hand side of (13), i.e. in the marginal benefit of prevention under cooperation. Such a term represents the *whole* positive impact that the costly prevention action has on the random premium (and, thus, on marginal utility of net consumption) in both states of the world. In other words and once again, under cooperation the positive impact of prevention on the premium is fully taken into account by policyholders.

Note that, despite full coverage, the representative policyholder faces some risks because his expected utility depends on his own realization of the illness. Indeed the latter changes the total number of health expenditures K given that $K=k+1$ if he gets the illness and $K=k$ if he is healthy.

[13] From an analytical point of view, full coverage is optimal since the Envelope Theorem can be fully applied when deriving the optimal coverage q under cooperation. See, in particular, the difference between equation (19) and (20) in Appendix 10.1 and 10.2 respectively.

6 Pool size and participating policies

We now compare the non-cooperative equilibrium and the cooperative outcome considering how policyholders' payoff changes with pool size. In particular we show that, as we expected, policyholders are better off under cooperation than in the non-cooperative equilibrium and that, under cooperation, their welfare is increasing in the pool size. Moreover, as the pool size goes to infinity, we show that the non-cooperative equilibrium converges to the second-best non-participating policy, whereas the cooperative outcome converges to the first-best allocation.

First of all consider that, when $n = 1$, the non-cooperative payoff and the cooperative outcome are obviously identical and depict a situation where the representative consumer is not insured:

$$EU_1 = pU\left(w - L(a)\right) + (1-p)\,U\left(w\right) - C(a)$$

The FOC with respect to prevention in this case is:

$$a_1^* : -pU'\left(w - L(a)\right)L'(a) = C'(a)$$

By definition, the optimal prevention level under cooperation satisfies $a^{*C} = \text{argmax}_a\, EU(a; n, q, a)$, whereas in the non-cooperative case prevention verifies $a^{*NC} = \arg\max_{a_i} EU(a_i; n, q, a_{-i}^{*NC})$. In other words: the cooperative outcome corresponds to a situation where agents optimize on both their own prevention level and on the prevention choices of others (provided they are the same) whereas in the non-cooperative equilibrium, prevention levels are constrained to be Nash-equilibrium prevention choices such that $a_i^{*NC} = a_{-i}^{*NC} = a^{*NC}$. Since the non-cooperative equilibrium can always be achieved cooperatively, we can state the following:

Remark 1 *Given a risk-sharing agreement with pool size strictly larger than* 1*, the cooperative outcome always dominates the non-cooperative equilibrium.*

Now let's consider how the outcome under cooperation changes with the pool size:

Lemma 2 *Under cooperation policyholders' expected utility is monotonically increasing with the size of the pool.*

Proof. See the Appendix 10.3. ■

Note that the previous result is an extension of Borch's rule (1962) to the case where consumers choose a self-insurance measure: the benefit of risk-sharing increases with the size of the pool.

Unfortunately, the proof of Lemma 2 cannot be easily extended to the non-cooperative case because of the partial coverage characterizing the non-cooperative equilibrium. Since the benefit of risk-sharing increases with the size of the pool in the non-cooperative equilibrium as well, we expect the policyholders' welfare to be increasing with the size of the pool also in the non-cooperative case. Our intuition is confirmed by the following simulations (see also Figure 1 below).

Simulations The reported results refer to CARA utility function[14]:

$$U(W) = -\frac{e^{-\rho W}}{\rho}$$

[14] Very similar results were obtained for other utility functions and are available from the authors on request.

where ρ represents the degree of constant absolute risk-aversion and was assigned the value of 0.5. Wealth w is 10 and the probability of illness p is 0.3. The cost of prevention is expressed by the function $C(a) = a^2$, while the loss caused by health expenditures is $L(a) = 1 - a$.

Policyholders' expected utility in the case of cooperation and in the case of non-cooperation are shown in Figure 1 below. First, as stated in Lemma 2, expected utility under cooperation is monotonically increasing in the size of the pool. Second, in line with Remark 1, the cooperative payoff dominates the non-cooperative equilibrium for every possible value of the pool size. Finally, as we expected, the non-cooperative payoff is (monotonically) increasing in the size of the pool.

Insert figure 1 about here

6.1 Asymptotic results

Remember that K is the total number of individuals in the pool experiencing the negative health shock. By the law of large numbers, as the number of policyholders tends towards infinity, the share of them suffering the illness $\left(\frac{K}{n}\right)$ tends towards the probability of illness p. This implies that, in a very large pool, the random premium with the participating policy $\frac{KqL(a)}{n}$ tends to the fixed premium $pqL(a)$ as in non-participating contracts (see Subsection 2.2 above), thus making the uncertainty on the premium disappear. This occurs both under the non-cooperative and under the cooperative strategy. Further, under cooperation, since full coverage is offered to policyholders, the fixed premium becomes $pL(a)$.

Let's consider the first-order condition for the prevention action in the non-cooperative equilibrium (10). As $n \to \infty$ the term $\frac{q}{n} \to 0$, while the term $\frac{k+1}{n} = \frac{K}{n} \to p$. Thus, the first-order condition becomes:

$$a_{\infty}^{*NC} : pU'\left(w - (1 - q + pq)\, L(a)\right)\left(-L'(a)(1-q)\right) = C'(a)$$

which is equivalent to (5). We can conclude that, as $n \to \infty$, the policyholders' payoff in the non-cooperative equilibrium is exactly the same as with the second-best contract with a fixed premium.

Let's now consider the optimal cooperative prevention action. As $n \to \infty$ the first-order condition (13) becomes:

$$a_{\infty}^{*C} : U'\left(w - pL(a)\right)\left(-pL'(a)\right) = C'(a) \tag{14}$$

By comparing (2) and (14) we can easily verify that:

$$a_{\infty}^{*C} \equiv a^{FB}$$

Thus, the consumers' payoff under cooperation when $n \to \infty$ replicates the first-best (3).

All the previous reasoning is stated in the following Lemma:

Lemma 3 *When the number of policyholders goes to infinity, a participating policy without cooperation tends to the second-best non-participating contract, while a participating policy in the case of cooperation converges to the first-best allocation.*

In what follows we verify that, contrary to the non-cooperative outcome, the cooperative choice of prevention is never an equilibrium in a one-shot game among policyholders because of the incentives to free-ride. We then study conditions such that cooperation can be sustained as equilibrium in a repeated game where policyholders interact an indefinite number of times.

7 Cooperation as an equilibrium

It can easily be shown that cooperation in the second stage of the game considered in Subsection 3.2 can never be an equilibrium in a static framework. Suppose that the optimal choice of q has already been taken by the pool in the previous stage ($q^{*C} = 1$). Under full coverage, the policyholder's best response a_i^{BR}, when all the other policyholders choose the cooperative strategy a^{*C}, is obtained as follows:

$$\max_{a_i^{BR}} EU_i = \sum_{k=0}^{n-1} b(k; n-1; p) \Big\{ pU \left[w - \frac{1}{n} \left(L(a_i^{BR}) + kL(a^{*C}) \right) \right] + (1-p)\, U \left[w - \frac{k}{n} L(a^{*C}) \right] \Big\} - C(a_i^{BR}) \tag{15}$$

Thus, the best response to the cooperative level of prevention a^{*C} is:

$$a_i^{BR}(n,1): \tag{16}$$
$$\sum_{k=0}^{n-1} b(k; n-1; p) \left\{ pU' \left[w - \tfrac{1}{n} \left(L(a_i^{BR}) + kL(a^{*C}) \right) \right] \left(-\tfrac{1}{n} L'(a_i^{BR}) \right) \right\} = C'(a_i^{BR})$$

which is lower than the cooperative prevention, i.e.:

$$a_i^{BR} < a^{*C}$$

The above result is not surprising: the costly prevention action performed by each policyholder in the pool exerts a positive externality on the random premium and the policyholder prefers to *free ride* on it. The most advantageous situation for a member of the pool is when all the other policyholders internalize the social benefit of prevention on the random premium, while he just internalizes the impact of his action on his own health expenditures in the event of illness. As is always the case for positive externalities, the market (or non-cooperative) solution implies that individuals' choices of prevention are inefficiently low. In our context this implies that cooperation cannot be sustained as an equilibrium in the one-shot interaction among policyholders.

In the following we will consider how the cooperative solution can be implemented. The issue is very similar to the cartel enforcement problem faced by firms in oligopolistic markets.

We investigate a case where the members of the pool interact for an uncertain number of periods. We will see that cooperation can be sustained as equilibrium if the members of the pool implement a punishment when they observe an insurance claim which is higher than expected. In fact, a high loss means that someone deviated from the cooperative choice by exerting a low level of prevention (see the last part of this section about extending the game to a stochastic loss function). As a consequence, to punish the deviator, the other policyholders exert the non-cooperative prevention level in all the subsequent periods.

In particular, we adapt the Folk Theorem to our environment by interpreting the time horizon as uncertain. Note that deviation by one policyholder is detected by the other members of the pool *only if* the deviator experiences the illness. Thus, we are in a stochastic environment regarding deviation observability.

Let's call δ the "probability-adjusted" discount factor.[15] We apply the Grim trigger strategy in our context: in the first period $t = 0$ the policyholder chooses the cooperative prevention level a^{*C}. In $t \geq 1$ the policyholder exerts a^{*C} if in each previous period any losses that have occurred are of size $L\left(a^{*C}\right)$, otherwise he exerts $a = a^{*NC}$ forever.

The Grim trigger strategy we have just described better fits a non-competitive environment, that is a situation where policyholders have no choice but to stay in the pool, as in the case of a monopolistic insurer.[16] We refer the reader to the conclusion of the paper for a discussion on how to make the punishment more coherent with a competitive health insurance market.

For expositional reasons, let's call $EU(a^{*C}, a^{*C}, n)$ the policyholder's payoff when he cooperates and $EU(a_i^{BR}, a^{*C}, n)$ his payoff when he deviates (see expression (15)). In the period after deviation, the policyholder will be detected only if the illness is realized, that is with probability p. In that case, all the other members of the pool will react by choosing the non-cooperative prevention level a^{*NC}.

Let's call $EU(a^{*NC}, a^{*NC}, n)$ the payoff the deviator obtains when he is detected. The strategy $a = a^{*NC}$ for the deviator is just the subgame perfect Grim trigger strategy in our game.

With probability $(1-p)$ the deviation is not detected and the policyholder obtains the payoff $EU(a_i^{BR}, a^{*C}, n)$ in the period after the deviation as well.[17] The previous reasoning is repeated in the subsequent periods. After a deviation the policyholder's payoff can be written as follows:

$$\begin{aligned}
&EU(a_i^{BR}, a^{*C}, n) + \delta p \left(\frac{1}{1-\delta} EU(a^{*NC}, a^{*NC}, n) \right) + \\
&\delta(1-p) \Bigg\{ EU(a_i^{BR}, a^{*C}, n) + \delta p \frac{1}{1-\delta} EU(a^{*NC}, a^{*NC}, n) + \\
&\quad \delta(1-p) \left[EU(a_i^{BR}, a^{*C}, n) + \delta p \frac{1}{1-\delta} EU(a^{*NC}, a^{*NC}, n) + \delta(1-p)[...] \right] \Bigg\}
\end{aligned}$$

Thus, the discounted payoff in the case of deviation is:

$$\frac{\delta p}{1-\delta} \sum_{t=0}^{\infty} \left(\delta\left(1-p\right)\right)^t EU(a^{*NC}, a^{*NC}, n) + \sum_{t=0}^{\infty} \left(\delta\left(1-p\right)\right)^t EU(a_i^{BR}, a^{*C}, n)$$

or:

$$\frac{1}{1-\delta\left(1-p\right)} \left[EU(a_i^{BR}, a^{*C}, n) + \frac{p\delta}{1-\delta} EU(a^{*NC}, a^{*NC}, n) \right];$$

[15] It is the product of the discount factor and the belief policyholders have regarding the probability that they will continue to interact from period to period.

[16] Note that, when no other insurers are active in the market, exclusion of the deviator from the pool would be a possible alternative punishment strategy. In that case the deviator, once detected, would obtain forever his expected utility when not insured: $EU_1 = pU\left(w - L(a)\right) + (1-p)\,U\left(w\right) - C(a)$. However, contrary to the non-cooperative equilibrium, exclusion of the deviator cannot be interpreted as an equilibrium of the one-shot game and, for this reason, we do not propose it as a punishment strategy.

[17] The strategy $a = a^{BR}$ for the deviator results from the Bellman principle:

$$\begin{aligned}
V_0 &= \max_a \left\{ EU(a, a^{*C}, n) + \delta\left[pV_N + (1-p)\,V_0\right] \right\} \\
&= EU(a^{BR}, a^{*C}, n) + \delta\left[pV_N + (1-p)\,V_0\right]
\end{aligned}$$

where V_0 is the present value of the policyholder's best reply when his deviation is not detected, whereas V_N is the present value of the policyholder's best response when his deviation is detected.

whereas if the policyholder cooperates forever, the discounted payoff he obtains is:

$$\sum_{t=0}^{\infty} \delta^t EU(a^{*C}, a^{*C}, n) = \frac{1}{1-\delta} EU(a^{*C}, a^{*C}, n).$$

From the previous reasoning, cooperation can be sustained as an equilibrium if the discounted payoff from cooperation dominates the discounted payoff from deviation, or when the inequality below holds:

$$\frac{1}{1-\delta} EU(a^{*C}, a^{*C}, n) \geq \frac{1}{1-\delta(1-p)} \left[EU(a_i^{BR}, a^{*C}, n) + \frac{p\delta}{1-\delta} EU(a^{*NC}, a^{*NC}, n) \right] \qquad (17)$$

From inequality (17) we can state the following:

Proposition 1 *The cooperative level of secondary prevention can be sustained as an equilibrium in a repeated game with uncertain horizon. When the cooperative equilibrium exists, it can be sustained only if the number of policyholders is not too high.*

Proof. See the Appendix 10.4. ■

The previous proposition states that only a pool size which is not too large is compatible with the cooperative equilibrium. In fact, incentives to free ride are higher in a large pool: when n is high, (i) by deviating from the cooperative prevention level the policyholder significantly decreases the cost of prevention, whereas deviation has almost no effect on the insurance premium and (ii) the punishment is less costly since the benefit from risk-sharing remains important. However, the cooperative equilibrium may not exist. As expected, existence is more likely the higher the policyholders' probability-adjusted discount factor δ.

Moreover, since the left and the right-hand side of (17) are equal for $n = 1$ (see the proof in Appendix 10.4), existence only occurs if the left-hand side of the inequality increases faster in the pool size than the right-hand side. In other words, cooperation can be obtained only if the benefit from risk-sharing for low size pools is larger under cooperation than under either non-cooperation or deviation.

In Figure 2 below we depict the left-hand side and the right-hand side of inequality (17) using the utility function and the parameter values considered before together with a probability-adjusted discount factor $\delta = 0.3$.[18] Simulations show that, in this example, the cooperative equilibrium exists for $n \leq 950$.[19] For a larger pool size, the equilibrium is non-cooperative and all policyholders choose the non-cooperative prevention level.

Insert Figure 2 about here

As basic comparative statics, we may take into account the effects of changing two crucial variables in the analysis of inequality (17): the risk aversion (relying on the parameter ρ) and the probability of illness (p), respectively. It is worth noting that an increase in the individuals' risk aversion from $\rho = 0.5$ to $\rho = 0.6$ leads to a large rise in the value of the pool size compatible with

[18] In Figure 2, the graphs depicting the left and the right-hand side of (17) are not represented for the pool size $n = 1$ because of the scale of the picture. At that point the two curves cross each other.

Also notice that, in the considered example, the combination of deviation and non-cooperation expressed in the right-hand side of inequality (17) is monotonically increasing in the pool size.

[19] Existence of the cooperative equilibrium can also be obtained with utility functions different from CARA.

cooperation from 950 to 3100. This result relies on the fact that, when risk-aversion increases, the benefit of risk-sharing among members in the pool is higher and utility from cooperation grows. Conversely, after a decrease in risk aversion from $\rho = 0.5$ to $\rho = 0.4$, we observe that the maximum pool size for which the cooperative equilibrium is sustainable is lower than 400.

Let's now consider the illness probability p. If it increases from $p = 0.3$ to $p = 0.6$, we observe that cooperation is compatible with a smaller pool size: the upper bound value of the pool size changes from 950 to 400. At first sight, this result may seem counterintuitive. In fact it relies on the effect the probability of accident has on the policyholders' optimal prevention level. Indeed, an increase in the probability of illness reduces the difference between the individuals' optimal prevention under cooperation and under non-cooperation (as it increases the benefit from exerting prevention) and thus decreases the differences between expected utilities. Symmetrically, reducing the probability from $p = 0.3$ to $p = 0.2$ increases the maximum size of the pool compatible with cooperation from 950 to 1600.

To sum up, when policyholders interact for an indefinite number of periods, the cooperative equilibrium can be achieved as long as $n \leq \hat{n}$, whereas only the non-cooperative equilibrium is sustainable if $n > \hat{n}$. Figure 3 describes the whole equilibrium schedule as a function of the pool size for the same numerical example we used above.

Insert Figure 3 about here

As a final remark, in our framework a deviation from the cooperative action can be inferred by observing an insurance claim larger than expected, provided that the policyholder who is deviating suffers the loss. As the literature on firms' collusion shows, it is possible to obtain cooperation also in settings where detection of cheating behaviors is even more uncertain than in our model (see Green and Porter 1984[20]). For this reason we guess that, also in our framework and under more stringent conditions, cooperation would still be sustained if the loss function was non-deterministic.

8 Policy considerations

We mentioned in the introduction that firms offering participating policies generally encourage long-run enrollment, prevention measures and all services promoting a healthy life (Broek at al., 2012).

We provide here some real examples of policy promoting long run enrollment. *CAMPA*, one of the most important health mutuals in the North of Italy, only offers three-year contracts, renewable every three years without any limit due to age or health conditions. It does not impose any age limit to subscription since individuals can enroll up to their 70th birthday and can still purchase a policy when they are older than 70, provided that a younger member of the family subscribes as well. Moreover, policyholders are explicitly encouraged to remain enrolled all life long. Finally, one of the missions of the mutual, is to promote preventive measures, check-ups and diagnostic screening: "We offer insurance against the financial risk associated with buying medical care and assure an appropriate and fast diagnostic screening and prevention".[21]

[20] They examine the nature of cartel self-enforcement in the presence of demand uncertainty. In particular, in their setting, demand fluctuations not directly observed by firms make the detection of deviation difficult to infer so that collusive equilibria are less likely and an unstable industry performance can occur. In fact, reversionary episodes, where price cut is performed by all firms in the cartel as a punishment strategy, can sometimes happen with no firm really defecting, simply because of low demand.

[21] See the home page of the website of CAMPA (consulted in January 2014): http://www.campa.it/index.php.

In 2006 AGF, a French mutual company, made an agreement with Danone such that its policyholders received a large reimbursement for the consumption of healthy food like "Danacol". AGF explained that this action was fully in line with its long term objective concerning policyholders' healthy behaviors and prevention (see Les Echos, 2006).

Differently from mutual firms, standard health insurers do not encourage long run enrollment. Contracts are strictly annual and their renewal is conditional on the policyholder's health conditions and age.[22,23]

In a normative perspective, to compare the efficiency of participating and non-participating policies we should take into account the best strategies that standard insurers could adopt if long term contracts were offered. In the concluding section, we provide some suggestions for a modeling strategy aimed at the comparison of the two policies.

However, let's take for a moment a positive perspective. Since standard health insurers do not offer long term contracts in the real world and do not generally use information on past behaviors to incentivate prevention, our results suggest that existing participating policies may be more effective in encouraging prevention than non-participating ones. To see this, suppose a cooperative equilibrium exists, as Figure 2 shows, for a size of the pool lower than *950*. For the same functions and parameter values used in the simulations of Section 7, we can evaluate expected utility derived from the second-best contract which represents a non-participating policy (see Subsection 2.2). Note that expected utility generated by a second-best policy with fixed premium is independent, by construction, of the size of the pool; whereas expected utility generated by a participating policy is increasing in the pool size, as Lemma 2 states. Simulations show that the participating policy dominates the second-best contract for a pool size larger than 250 (see Figure 4).[24] This implies that, in this example, a large range of pool sizes exists such that cooperation is sustained as an equilibrium and the participating policy is more efficient than the second-best contract.

Insert Figure 4 about here

As a last remark, our results echo a recent debate on insurance regulation and mutual companies. In 2009 an association of French mutual insurance companies (ROAM) started to protest against the new measures that "Solvency 2" was going to introduce to increase the security of consumers; the protest spread to the entire European insurance sector one year later (see http://www.stopsolvency2.com/). The European Insurance and Reinsurance Federation (CEA) claimed that many players would not be able to continue doing business and that a sharp fall in supply may occur. Indeed, because of the administrative cost introduced by the new regulation, the latter would be threatening for small or middle-sized organizations. As a result, in many cases, insurance prices would be pushed higher without improving the security of consumers. In this respect, our results suggest that the survival of small and middle-sized firms offering participating policies is also important for efficiency reasons, so that possible drawbacks of the new regulation could be greater than alleged.

[22] The Affordable Care Act approved recently in the US has many objectives; one of them is to protect policyholders from the insurers' practice of refusing policy renewal in the case of serious health conditions.

[23] Interestingly, long term contracts are instead offered by stock insurers in other markets. For example front-loaded contracts in life insurance generate a partial lock-in of consumers: contracts that are more front-loaded have a lower present value of premiums over the period of coverage (see Hendel and Lizzeri 2003).

[24] These simulations are available upon request to the authors.

9 Conclusion

Previous literature showed that a copayment on expenditures for preventive care is necessary to encourage prevention choices since, with standard contracts, policyholders do not internalize the positive impact of their action on the insurance premium.

In this paper we analyzed an alternative type of contract, the prevention choices it leads to and its welfare properties. We showed that participating policies allow policyholders to make non-myopic choices concerning preventive care. More specifically, members of risk-sharing arrangements may find it convenient to choose a cooperative prevention strategy and fully cope with the appropriate level of prevention. A necessary condition is that policyholders enroll for a sufficient amount of time, so that cooperation becomes a worthy strategy. If the size of the pool is not too large, the upper bound of cooperation results from the trade-off between mutualization and free-riding issues. This is consistent with the general view that small pools allow for better monitoring of policyholders' behaviors and with the empirical evidence that firms offering health participating policies are typically small (see Caire 2009, Kerleau 2009 and Lippi Bruni *et al.* 2012).

The cooperative behavior we consider in the model does not require an individual to be empathetic with other members within the pool, but originates from an absolutely standard utility maximizer attitude. However, supportive, fair and conditional cooperative behaviors can be considered plausible in a mutual agreement, given the very specific nature of the participating contract. In other words, willingness to cooperate may be higher for individuals who self-select in organizations selling participating policies. For example, solidarity principles explicitly mentioned in all mutual insurance articles of association/incorporation, if shared by the policyholders, could well provide a better route to cooperation[25]. Moreover, cheating behaviors can also entail some psychological costs to the deviator in terms of, for example, lower self-esteem or a social stigma. Obviously, some kind of prosocial attitude or altruism may push towards cooperation as well (see Alger and Weibull, 2010, on the impact of altruism on risk-sharing).

Although remaining in a framework where individuals are characterized by purely selfish preferences, members of the mutual agreement partially know each other since they assemble for periodical meetings. This suggests that some partial peer monitoring is possible, letting cooperative equilibria easier to be sustained.

As a final consideration, cooperation has been analyzed under a punishment strategy that better fits a non-competitive setting (we assumed that, if a deviation is detected, all policyholders turn to their non-cooperative strategy). It might therefore be interesting to analyze to what extent our results can be generalized to other punishment strategies, compatible with a competitive environment. However, such analysis depends largely on the kind of policies (participating vs. non-participating) offered by the competitors and on how the latter make use of historical data when offering non-participating policies (see for example Moreno *et al.*, 2006 on bonus-malus schemes used in car insurance). Interestingly, this framework would allow a comparison of the efficiency properties of participating and non-participating policies when prevention matters and non-participating policies promote policyholders long run enrollment too; this question is left to future research.

[25] The beneficial matching between agents characterized by a similar "mission" (or social attitude) has been analyzed by Besley and Ghatack (2005), where moral hazard can be solved more cheaply if employer and employee share the same motivation.

10 Appendix

10.1 Proof of Lemma 1

We solve the game in two stages. In the first stage, policyholders "cooperatively" choose the level of coverage q and, in the second one, each individual chooses non-cooperatively his level of prevention a. Remember that:

$$EU_i^{NC}(q, a_i, a_{-i}) = \sum_{k=0}^{n-1} b(k; n-1; p)\{pU\left(W_L^{NC}\right) + (1-p)\,U\left(W_0^{NC}\right)\} - C(a_i)$$

where $W_L^{NC} = w - \frac{q}{n}\left(L(a_i) + kL(a_{-i})\right) - L(a_i) + qL(a_i)$ and $W_0^{NC} = w - \frac{kq}{n}L(a_{-i})$.

In the second stage each policyholder maximizes his utility w.r.t. a_i. The symmetric Nash equilibrium, $a^{NC}(q)$, thus solves

$$\left(1 - q + \frac{q}{n}\right)\sum_{k=0}^{n-1} b(k; n-1; p)pU'\left[w - \left(1 - q + \frac{q}{n}(k+1)\right)L(a)\right] + \frac{C'(a)}{L'(a)} = 0 \qquad (18)$$

while, in the first stage, the optimal coverage solves

$$\max_q V(q) \equiv EU_i^{NC}(q, a^{NC}(q), a^{NC}(q)).$$

since, according to the envelope theorem, $\frac{\partial EU_i^{NC}}{\partial a_i}(q, a^{NC}(q), a^{NC}(q)) = 0$, the optimal coverage q is implicitly defined as:

$$V'(q) = \frac{\partial EU_i^{NC}}{\partial q}(q, a^{NC}(q), a^{NC}(q)) + \frac{\partial EU_i^{NC}}{\partial a_{-i}}(q, a^{NC}(q), a^{NC}(q))\frac{\partial a^{NC}}{\partial q}(q) = 0 \qquad (19)$$

The following part of the proof is in two steps. In the first step we prove that the first term in the previous equation is equal to zero in the case of full coverage. In the second step we prove that the second term of the previous equation is negative in the case of full coverage. This implies that full coverage is not the optimal coverage and that the optimal coverage is lower than 1.

Step 1. Let's first prove that the first term in (19) is nill for $q = 1$

$$\begin{aligned}
\frac{\partial EU_i^{NC}}{\partial q}(q, a^{NC}(q), a^{NC}(q)) &= L(a^{NC}(q)) \textstyle\sum_{k=0}^{n-1} b(k; n-1; p)\{pU'\left[W_L^{NC}\right]\frac{n-(k+1)}{n} - (1-p)U'\left[W_0^{NC}\right]\frac{k}{n}\} \\
&= pL(a^{NC}(q)) \textstyle\sum_{k=0}^{n-1} b(k; n-1; p)U'\left[W_L^{NC}\right]\frac{n-(k+1)}{n} \\
&- (1-p)L(a^{NC}(q)) \textstyle\sum_{k=0}^{n-1} b(k; n-1; p)U'\left[W_0^{NC}\right]\frac{k}{n}
\end{aligned}$$

where the first term equals 0 for $k = n-1$ and the second term equals 0 for $k = 0$, thus:

$$\begin{aligned}
\frac{\partial EU_i^{NC}}{\partial q}(q, a^{NC}(q), a^{NC}(q)) &= pL(a^{NC}(q)) \sum_{k=0}^{n-2} b(k; n-1; p) U' \left[W_L^{NC}\right] \tfrac{n-(k+1)}{n} \\
&- (1-p) L(a^{NC}(q)) \sum_{k=1}^{n-1} b(k; n-1; p) U' \left[W_0^{NC}\right] \tfrac{k}{n} \\
&= pL(a^{NC}(q)) \sum_{k=0}^{n-2} \tfrac{(n-1)!}{k!(n-1-k)!} p^k (1-p)^{n-1-k} U' \left[W_L^{NC}\right] \tfrac{n-(k+1)}{n} \\
&- (1-p) L(a^{NC}(q)) \sum_{k=1}^{n-1} \tfrac{(n-1)!}{k!(n-1-k)!} p^k (1-p)^{n-1-k} U' \left[W_0^{NC}\right] \tfrac{k}{n} \\
&= \tfrac{L(a^{NC}(q))}{n} \sum_{k=0}^{n-2} \tfrac{(n-1)!}{k!(n-1-(k+1))!} p^{k+1} (1-p)^{n-1-k} U' \left[W_L^{NC}\right] \\
&- \tfrac{L(a^{NC(q)})}{n} \sum_{k=1}^{n-1} \tfrac{(n-1)!}{(k-1)!(n-1-k)!} p^k (1-p)^{n-k} U' \left[W_0^{NC}\right] \\
&= \tfrac{L(a^{NC}(q))}{n} \sum_{k=0}^{n-2} \tfrac{(n-1)!}{k!(n-1-(k+1))!} p^{k+1} (1-p)^{n-1-k} U' \left[w - \left(\tfrac{q(k+1)}{n} + 1 - q\right)\right] \\
&- \tfrac{L(a^{NC}(q))}{n} \sum_{k=1}^{n-1} \tfrac{(n-1)!}{(k-1)!(n-1-k)!} p^k (1-p)^{n-k} U' \left[w - \tfrac{qk}{n}\right]
\end{aligned}$$

A change of index with respect to the binomial probability on the left side gives:

$$\begin{aligned}
\frac{\partial EU_i^{NC}}{\partial q}(q, a^{NC}(q), a^{NC}(q)) &= \tfrac{L(a^{NC}(q))}{n} \sum_{k=1}^{n-1} \tfrac{(n-1)!}{(k-1)!(n-1-k))!} p^k (1-p)^{n-k} U' \left[w - \left(\tfrac{qk}{n} + 1 - q\right))\right] \\
&- \tfrac{L(a^{NC}(q))}{n} \sum_{k=1}^{n-1} \tfrac{(n-1)!}{(k-1)!(n-1-k)!} p^k (1-p)^{n-k} U' \left[w - \tfrac{qk}{n}\right] \\
&= \tfrac{L(a^{NC}(q))}{n} \sum_{k=1}^{n-1} \tfrac{(n-1)!}{(k-1)!(n-1-k)!} p^k (1-p)^{n-k} \Big\{ U' \left[w - \left(\tfrac{qk}{n} + 1 - q\right)\right] \\
&- U' \left[w - \tfrac{qk}{n}\right] \Big\}
\end{aligned}$$

Therefore when $q = 1$ the first term of (19) is $\frac{\partial EU_i^{NC}}{\partial q}(1, a^{NC}(1), a^{NC}(1)) = 0$ such that for $q = 1$ (19) can be rewritten as:

$$V'(1) = \frac{\partial EU_i^{NC}}{\partial a_{-i}}(1, a^{NC}(1), a^{NC}(1)) \frac{\partial a^{NC}}{\partial q}(1)$$

Step 2. Now we prove that $V'(1) < 0$, which implies that $q = 1$ is not the optimal coverage. First,

$$\frac{\partial EU_i^{NC}}{\partial a_{-i}}(q, a^{NC}(q), a^{NC}(q)) = -\frac{q}{n} L'^{NC}(q)) \sum_{k=0}^{n-1} b(k; n-1; p) k \left[pU'\left(W_L^{NC}\right) + (1-p) U'\left(W_0^{NC}\right)\right]$$

and is positive for all levels of q.
For $q = 1$ and given that K is the total number of losses in the pool of insurees, the previous expression equals:

$$\frac{\partial EU_i^{NC}}{\partial a_{-i}}(1, a^{NC}(1), a^{NC}(1)) = -\frac{1}{n} L'^{NC}(1)) \sum_{K=0}^{n} b(K; n; p)(K-1) U' \left(w - \frac{K}{n} L(a^{NC}(1))\right) > 0$$

Now, writing (18) as $f(a,q)=0$, and fully differentiating with respect to a and q we have $\frac{\partial a^{NC}}{\partial q} = -\frac{\partial f(a,q)}{\partial q} / \frac{\partial f(a,q)}{\partial a}$. In particular, we have:

$$\frac{\partial f}{\partial q}(a,q) = \sum_{k=0}^{n-1} b(k;n-1;p)p\left[\left(1-q+\frac{q}{n}\right)\left(1-\frac{k+1}{n}\right)L(a)U^{''}(W_L^{NC}) - \frac{n-1}{n}U'(W_L^{NC})\right] < 0$$

and

$$\frac{\partial f}{\partial a}(a,q) = -\left(1-q+\tfrac{q}{n}\right)\textstyle\sum_{k=0}^{n-1} b(k;n-1;p)p\left(1-q+\frac{q}{n}(k+1)\right)L^{''}\left(W_L^{NC}\right)+\frac{C^{''}(a)L'(a)-C'^{''}(a)}{[L'^2} < 0$$

Therefore, for all levels of coverage $\frac{\partial a^{NC}}{\partial q} < 0$. In particular, for $q=1$, we have:

$$\frac{\partial a^{NC}}{\partial q}(1) = \frac{\sum_{k=0}^{n-1} b(k;n-1;p)p\left[\frac{1}{n}\left(1-\frac{k+1}{n}\right)L(a)U^{''}\left(w-\frac{k+1}{n}L(a^{NC}(1))\right)-\frac{n-1}{n}U'\left(w-\frac{k+1}{n}L(a^{NC}(1))\right)\right]}{\frac{1}{n}\sum_{k=0}^{n-1} b(k;n-1;p)p\frac{1}{n}(k+1)L^{''}\left(w-\frac{k+1}{n}L(a^{NC}(1))\right)-\frac{C^{''}(a)L'(a)-C'^{''}(a)}{[L'^2}} < 0$$

10.2 Proof of Corollary 1

Under cooperation and in the first stage, the optimal coverage solves

$$\max_q V(q) \equiv EU^C(q, a^C(q), a^C(q)).$$

applying the envelope theorem:

$$V'(q) = \frac{\partial EU^C}{\partial q}(q, a^C(q), a^C(q)) = 0 \tag{20}$$

Since $\frac{\partial EU^C}{\partial q}(q, a^C(q), a^C(q)) = \frac{\partial EU_i^{NC}}{\partial q}(q, a^{NC}(q), a^{NC}(q))$, we can now refer to Step 1 in the proof of Lemma 1 to state that (20) is verified for $q=1$.

10.3 Proof of Lemma 2

Let's first define $\tilde{x}_i$ as the stochastic wealth of individual i given that k agents among the $n-1$ others fall ill. Full coverage implies that, under cooperation $\tilde{x}_i = w - \tilde{L}_i$ where $L_i = \frac{k+1}{n}L(a)$ with probability p and $\frac{k}{n}L(a)$ with probability $(1-p)$.

$EU(a^{*C}, a^{*C}, n)$ is then increasing in n if $EU(a^{*C}, a^{*C}, n+1)$ is less risky than $EU(a^{*C}, a^{*C}, n)$, that is if adding a new member to the pool decreases the aggregate risk. This can be written as $\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}$ less risky than $\sum_{i=1}^{n-1}\frac{\tilde{x}_i}{n-1}$ or

$$\sum_{i=1}^{n}\frac{\tilde{x}_i}{n} + \tilde{\varepsilon} = \sum_{i=1}^{n-1}\frac{\tilde{x}_i}{n-1} \tag{21}$$

with $E\left[\tilde{\varepsilon}\,\middle|\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}\right]=0$. Equation (21) can be rewritten as:

$$\tilde{\varepsilon} = \frac{n\sum_{i=1}^{n-1}\tilde{x}_i - (n-1)\sum_{i=1}^{n}\tilde{x}_i}{n(n-1)} = \frac{\sum_{i=1}^{n}\tilde{x}_i - n\tilde{x}_n}{n(n-1)}$$

Therefore,

$$E\left[\tilde{\varepsilon}\left|\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}\right.\right]=0 \text{ if and only if } \frac{\sum_{i=1}^{n}\tilde{x}_i}{n\left(n-1\right)}=\frac{1}{n-1}E\left[\tilde{x}_n\left|\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}\right.\right]$$

that is if and only if:

$$E\left[\tilde{x}_n\left|\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}\right.\right]=\frac{\sum_{i=1}^{n}\tilde{x}_i}{n}$$

Now, as the $\tilde{x}_i$ are i.i.d., we have:

$$E\left[\tilde{x}_n\left|\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}\right.\right]=E\left[\tilde{x}_k\left|\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}\right.\right]\forall k$$

and

$$E\left[\tilde{x}_n\left|\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}\right.\right]=\frac{\sum_{k=1}^{n}E\left[\tilde{x}_k\left|\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}\right.\right]}{n}=E\left[\frac{\sum_{k=1}^{n}\tilde{x}_k}{n}\left|\sum_{i=1}^{n}\frac{\tilde{x}_i}{n}\right.\right]=\frac{\sum_{k=1}^{n}\tilde{x}_k}{n}$$

10.4 Proof of Proposition 1

We show that the left-hand side of (17) is equal to its right-hand side for $n=1$; whereas the left-hand side is lower than the right-hand side for $n=\infty$. Remember that, from Lemma 2, $EU(a^{*C},a^{*C},1)$ is monotonically increasing in n. Thus, a cooperative equilibrium exists if the left-hand side of (17) crosses its right-hand side from above, implying that the size of the pool must be sufficiently low.

- For $n=1$, the sole policyholder will always choose the optimal prevention level. Thus, $EU(a^{*C},a^{*C},1)=EU(a^{*NC},a^{*NC},1)=EU(a_i^{BR},a^{*C},1)$ and condition (17) holds with equality.
- When $n\rightarrow\infty$ the impact of one policyholder's prevention on the premium is negligible so that deviation is always profitable. This can be seen by rewriting inequality (17) with $n\rightarrow\infty$

$$\begin{gathered}\frac{1}{1-\delta}\left[U(w-pL\left(a^{*C}\right))-C\left(a^{*C}\right)\right]\geq \\ \frac{1}{1-\delta\left(1-p\right)}\left[U(w-pL\left(a^{*C}\right))+\frac{p\delta}{1-\delta}EU(a^{*NC},a^{*NC},\infty)\right]\end{gathered}\tag{22}$$

where

$$EU(a^{*NC},a^{*NC},\infty)=pU\left[w-(1-q+pq)L(a^{*SB})\right]+(1-p)U\left[w-pqL(a^{*SB})\right]-C(a^{*SB})$$

Rearranging we can write:

$$C\left(a^{*C}\right)\leq p\delta\left[EU(a^{*NC},a^{*NC},\infty)-\frac{U(w-pL\left(a^{*C}\right))}{1-\delta\left(1-p\right)}\right]$$

Since $\frac{1}{1-\delta(1-p)}>1$ and $EU(a^{*NC},a^{*NC},\infty)<U(w-pL\left(a^{*C}\right))$, the right-hand side of the previous inequality is negative so that the latter is never satisfied. Thus, deviation is always profitable for $n\rightarrow\infty$ and (17) does not hold.

References

[1] Alger, I. and W. Weibull (2010), "Kinship, Incentives, and Evolution", *The American Economic Review* 100, 1725-1758.

[2] Barigozzi F. (2004), "Reimbursing Preventive Care", *Geneva Papers on Risk and Insurance Theory*, 29(2), 165-186.

[3] Besley, T.J. (1988), "Optimal Reimbursement Health Insurance and the Theory of Ramsey Taxation", *Journal of Health Economics* 7, 321–336.

[4] Besley T. and M. Ghatak (2005), "Competition and Incentives with Motivated Agents", *The American Economic Review* 95(3), 616-636.

[5] Bramoullé, Y. and Kranton R. (2007), "Risk Sharing Across Communities", *The American Economic Review* 97(2), 70-74.

[6] Borch, K. (1962), "Equilibrium in a reinsurance market", *Econometrica*, 30, 424–444.

[7] Broek S., Buiskool B.J., Vennekens A. and van der Horst R. (2012), "Study on the current situation and prospects of mutuals in Europe", Final report by European Commission, Project number: BA03954

[8] Caire, G. (2009), "La nouvelle logique d'action des mutuelles de santé : la mise en place de priorité santé mutualiste", *Recma, Revue Internationale d'Economie Sociale*, 311, 22–35.

[9] Cummins, JD, M Rubio-Misas, and H Zi, (2004), "The Effect of Organizational Structure on Efficiency: Evidence from the Spanish Insurance Industry", *Journal of Banking & Finance* 28, 3113–3150.

[10] Cummins, JD, MA Weiss, and H Zi (1999), "Organizational Form and Efficiency: The Coexistence of Stock and Mutual Property-Liability Insurers", *Management Science* 45, 1254–1269.

[11] Cummins, JD and H Zi (1998) ,"Comparison of Frontier Efficiency Methods: An Application to the U.S. Life Insurance Industry", *Journal of Productivity Analysis* 10, 131–152.

[12] Cutler, D.M. and R.J. Zeckhauser (2000), "The anatomy of health insurance", in Culyer, A., Newhouse, P. (Eds.), Handbook of Health Economics. North Holland.

[13] Doherty, N. and Dionne, D. (1993), "Insurance with undiversifiable risk: Contract structure and organizational form of insurance firms", *Journal of Risk and Uncertainty*, 6, 187-203.

[14] Ehrlich, I. and Becker, G. (1972), "Market insurance, self-insurance and self-protection", *Journal of Political Economy*, 80, 623-648.

[15] Ellis R. and Manning W. (2007), "Optimal health insurance for prevention and treatment", *Journal of Health Economics*, 26, 1128–1150

[16] Fafchamps, M. and Lund, S. (2003), "Risk Sharing Networks in Rural Philippines", *Journal of Development Economics*, 71(2), 261-287.

[17] Friesen G. (2007), "On the Lack of Participating Policy Usage by Stock Insurance Companies", *Journal of Risk and Insurance*, 74(1), 185-206.

[18] Genicot, G. and D. Ray (2003), "Group Formation in Risk-Sharing Arrangements", *Review of Economic Studies* 70, 87-113.

[19] Geoffard, P.-Y., (2000), "Assurance Maladie: la Gestion du Risque Long", *Revue d'Économie Politique,* 110(4), pp. 457-482.

[20] ICMIF (International Cooperative and Mutual Federation), Global Mutual Market Share report 2010, available at http://www.icmif.org/mms2010.

[21] Hendel, I. and A. Lizzeri, (2003). "The Role Of Commitment In Dynamic Contracts: Evidence From Life Insurance", The Quarterly Journal of Economics, 118(1), 299-327.

[22] Kenkel, D.S. (2000), "Prevention," in Handbook of Health Economics, Culyer and Newhause (Eds.), 1675–1719.

[23] Kerleau, M. (2009), "Le nouveau cadre institutionnel de la protection sociale complémentaire d'entreprise: quels enjeux pour les mutuelles santé?", *Recma*, *Revue internationale d'*écon*omie sociale*, 312, 15–34.

[24] Krishnaswami, S., and Pottier, S. (2001), "Agency Theory and Participating Policy Usage: Evidence From Stock Life Insurers", *Journal of Risk and Insurance*, 68(4), 659-683.

[25] Lee, W. and Ligon, J. (2001), "Moral Hazard in risk pooling arrangements", *The Journal of Risk and Insurance*, 68(1), 175-190.

[26] Les Echos N°19641 (2006), "AGF et Danone s'associent contre le cholestérol". Available at http://www.lesechos.fr/05/04/2006/LesEchos/19641-206-ECH_agf-et-danone-s-associent-contre-le-cholesterol.htm

[27] Ligon, J. and Thistle, P. (2005), "The formation of mutual insurers in markets with adverse selection", *Journal of Business*, 78, 529-555.

[28] Lippi Bruni, M., Rago S. and Ugolini C. (2012), "Il ruolo delle mutue sanitarie integrative", Società Il Mulino, ISBN 978-88-15-24100-9

[29] Mayers, D. and Smith, C. (1986), "Ownership Structure and Control: The Mutualization of Stock Life Insurance Companies", *Journal of Financial Economics*, 16, 73-98.

[30] Moreno I., Velazquez F. and Watt R. (2006), "Can Bonus-Malus Alleviate Insurance Fraud?", *The Journal of Risk and Insurance*, 73(1), 123-151.

[31] Picard, P. (2009), "Participating insurance contracts and the Rothschild-Stiglitz equilibrium puzzle," working paper Ecole Polytechnique, 2009-2030.

[32] Smith, B. and Stutzer, M. (1990), "Adverse Selection, aggregate uncertainty and the role for mutual insurance contracts", *Journal of Business*, 63, 493–510.

[33] Smith, B. and Stutzer, M. (1995), "A Theory of Mutual Formation and Moral Hazard with Evidence from the History of the Insurance Industry", *Review of Financial Studies*, 8, 545-577.

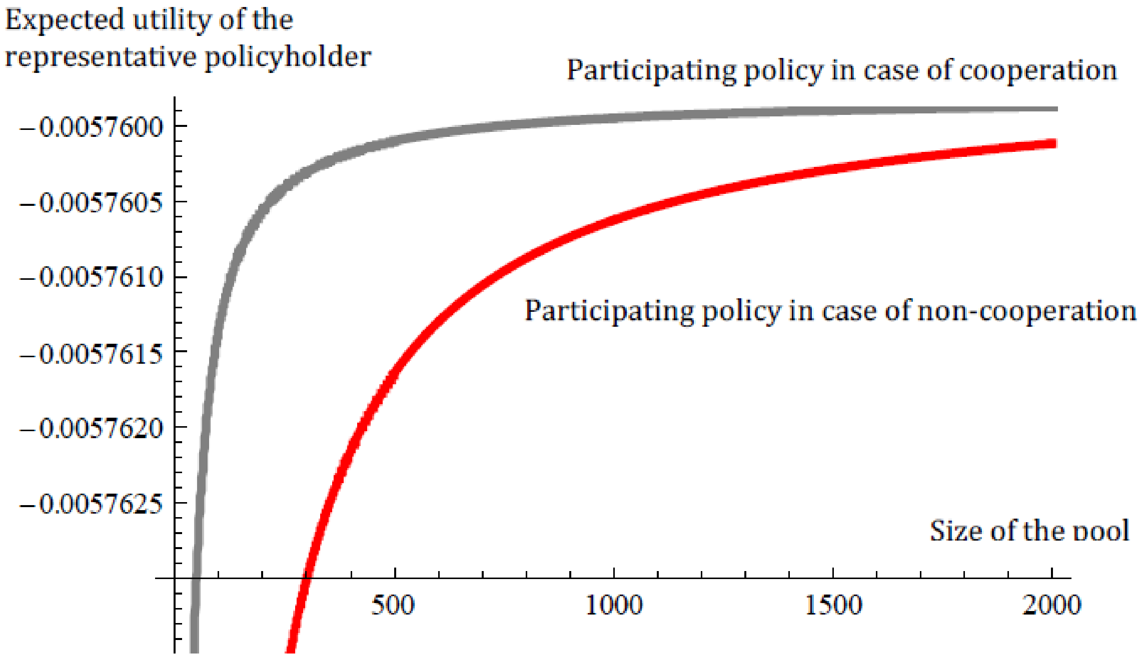

Figure 1: Simulation shows that even the non-cooperative equilibrium curve is monotonically increasing in the pool size.

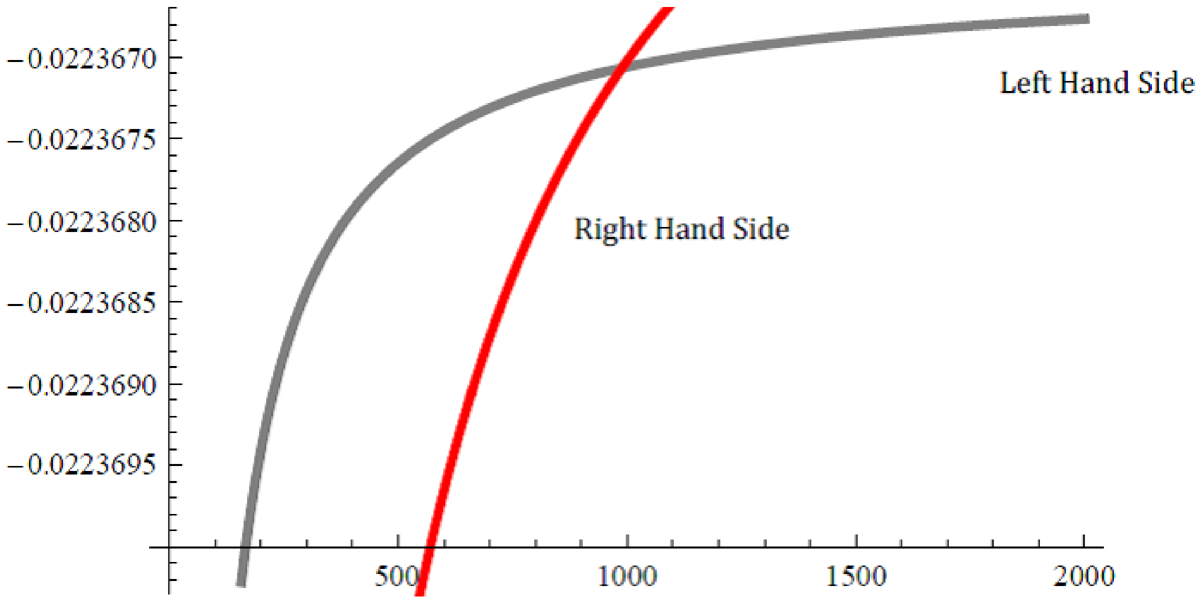

Figure 2: the curve representing the left hand side of equation (23) dominates the curve representing the right hand side until $\hat{n} \cong 950$.

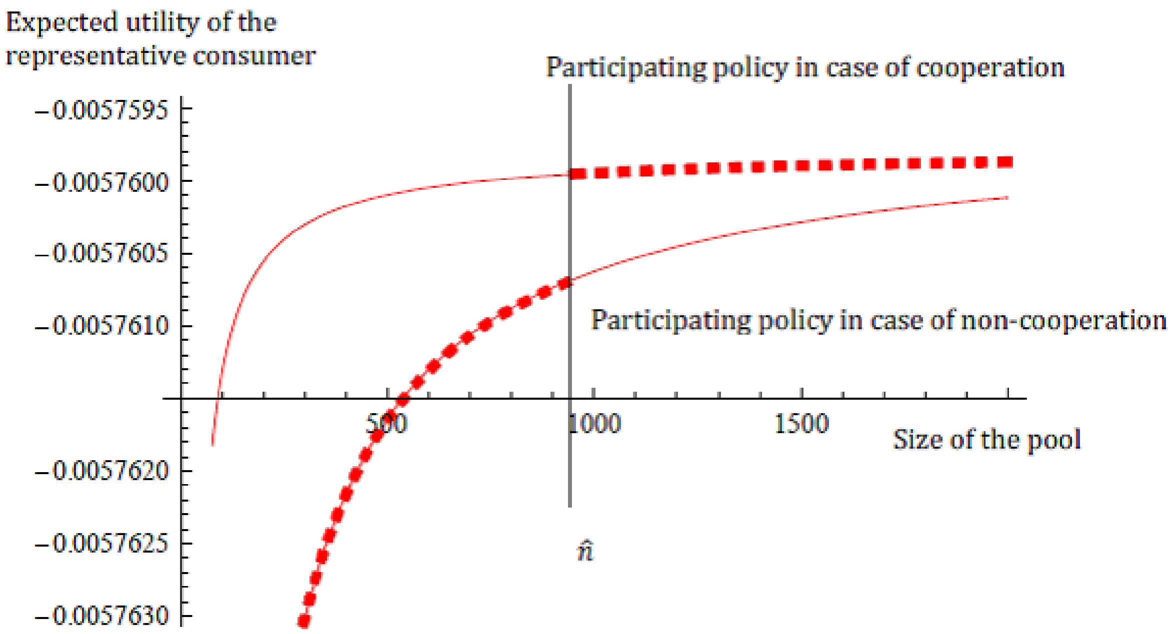

Figure 3: Solid lines describe the whole equilibrium schedule as a function of the pool size in the case where the cooperative equilibrium exists.

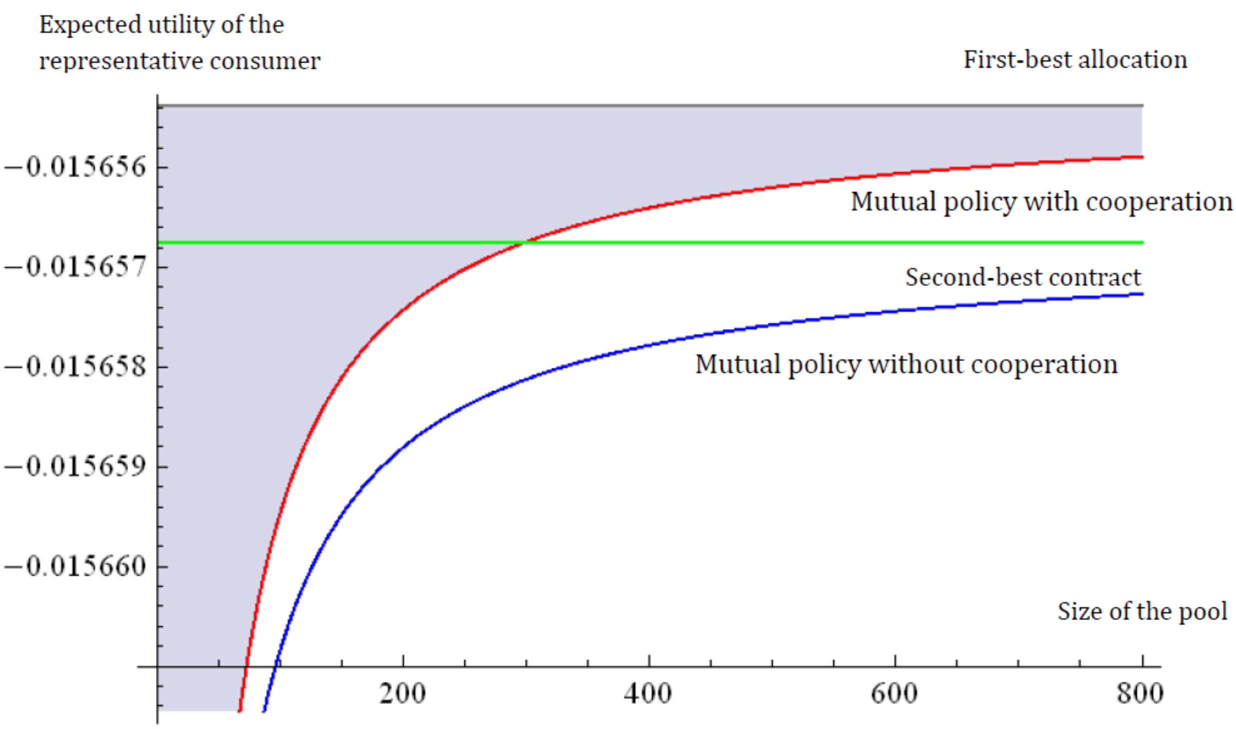

Figure 4: First-best, second-best, participating policies with and without cooperation with the CARA function.

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
http://www.dse.unibo.it