The Use of Risk Adjustment and Risk Sharing in Reducing the Welfare Loss produced by Health Plan Selection and Inefficiency

Manuel García Goñi
Universidad Complutense de Madrid

November 8, 2009

Abstract

Health care cost escalation is a serious problem in many countries and many researchers point to managed care through capitation as an important tool for controlling costs while fostering cost-effectiveness. While capitation can create desirable efficiency incentives, it also creates strong selection incentives. This paper compares the effectiveness of risk adjustment and risk sharing strategies through different reimbursement payment systems for reducing the welfare loss due to selection in the health care market. Selection and efficiency incentives enter in a three-stage model in which consumers choose provider, profit maximizing plans decide the schedule of services offered, and regulator select the payment system that minimizes a social welfare loss. Minimum welfare loss risk adjustment is superior to other risk adjustment strategies, but only uniformly superior to risk sharing when the quality of the information used by the payer is high enough.

Acknowledgment: I thank Pedro Albarrán, Randy Ellis, Kevin Lang, Matilde Machado, Tom McGuire, and participants at seminar at Boston University, and at ATINER, AES, iHEA and ESWC Congresses for their helpful comments. This research started with the financial support from the Centers for Medicare & Medicaid Services, award number 30-P-91713/1-01 that is gratefully acknowledged, and ended with the support by an unrestricted educational grant awarded jointly to the Universities Carlos III de Madrid and Pompeu Fabra de Barcelona by The Merck Foundation, the philanthropic arm of Merck Co. Inc., White House Station, New Jersey, USA. The remaining errors are the author's responsibility.
JEL: I100, I110, H510
Keywords: Managed Care, selection, risk adjustment, risk sharing.
Address for correspondence: Manuel García Goñi, Universidad Complutense de Madrid, Departamento de Economía Aplicada II. Campus de Somosaguas. 28223 Pozuelo de Alarcón (Madrid). Email: mg-goni@ccee.ucm.es Tel +34913942455 Fax +34913942457.

1 Introduction

Health care cost escalation is a serious problem in many countries and many researchers and policy-makers point to managed care as an important tool for controlling costs while fostering cost-effectiveness. Managed care is now the dominant form of insurance coverage in the U.S. and is being promoted and considered carefully in many other countries. For managed care plans to have appropriate incentives to contain costs, it is often recommended that plans receive a capitated payment so that they can benefit from any cost savings achieved. While capitation can create desirable efficiency incentives - plans want to manage care effectively - it also creates strong selection incentives: avoiding unprofitable enrollees may be easier than managing care (Newhouse, 1996). This paper addresses this central tradeoff between efficiency and selection incentives and attempts to understand how well various strategies can be used to improve the tradeoff.

The traditional solution to the selection problem is risk adjustment, in which the regulator pays premiums to health plans based on cost weights attached to variables that predict future health costs for each individual. Risk adjustment is commonly used in order to improve the predictability of future total health care costs but also for pharmaceutical costs (Zhao et al., 2005; García-Goñi et al., 2009). While conventional risk adjustment models currently in use are an improvement over simply using demographic information (age and gender) to predict costs, it has been shown (Chapman, 1997; Shen and Ellis, 2002a) that they remain deficient in that health plans still have incentives to use their own private information to better predict the cost of each insured and to select profitable enrollees (good risks) if they are able to.

There exists a number of research providing evidence of risk selection under managed care.[1] However, it is unclear how they may select. Since some programs as Medicare require open enrollment, plans cannot directly turn down unprofitable enrollees ("bad risks"). Instead, health plans may affect selection indirectly, as reviewed in Van de Ven and Ellis (2000). These indirect strategies include referring patients with serious chronic conditions to providers in a different health plan, or providing poor quality of care to identified bad risk individuals, both of which could be considered "dumping" strategies, based on individual characteristics (Ellis, 1998; Shen and Ellis, 2002b). Another set of strategies include structuring coverage in such a way that is unattractive for bad risks, by not contracting with physicians who have the best reputation of treating patients with

[1] Brown et al. (1993) uses pre-enrollment costs and use of services to show that enrollees in the HMOs are healthier than those in FFS health plans. Litchenstein et al. (1991) and Kravitz et al. (1992) use self-reported health status and medical conditions and obtain the same result. Other studies, as Riley et al. (1991), compare the mortality rate and get the conclusion that it is lower in the HMOs than in FFS plans. A last example is Riley et al. (1996), that uses different health status measures obtained from the MCBS data for year 1994 and also obtain that Medicare HMO enrollees are healthier than beneficiaries in FFS.

chronic illnesses, or by underproviding quality or quantity of specific services which tend to attract bad risks. These "skimping" strategies all imply service distortion, and have been analyzed in Glazer and McGuire (2000) and Frank, Glazer and McGuire (2000). In those papers, health plans use quality of services as the tool for selection, and have incentives to overprovide quality in some services and underprovide the quality in others. As Glazer and McGuire note, health plans trying not to attract to bad risks does not mean that health plans have incentives to decrease the quality of all services offered. Quality is still desirable if it attracts the good risks. Glazer and McGuire (2002a) use only prospective information but not in a regression-based approach. Instead, Glazer and McGuire choose optimal payment weights so as to maximize efficiency in offered quality of services. With respect to empirical evidence on the type of selection, Cao and McGuire (2003) and Cao (2003) find evidence of service distortion in Medicare where HMOs ration services differently, providing more health care in primary care services while less in mental health care services. In contrast, Ellis and García-Goñi (2009) find that there is no evidence of selection through dumping strategies studying the relationship between the HMO market shares and the different payment rates for aged and disabled in the Medicare market. These results point to skimping strategies as the explanation for the selective behavior.

Risk sharing is an alternative to risk adjustment. Similar to using retrospective information for risk adjustment, risk sharing also uses *ex post* information for payments. However, the essence of risk sharing is that costs, rather than other indicators of need, are used to make payments. Ellis and McGuire (1986) introduced the notion of supply-side cost sharing and coined the term "mixed system" for paying hospitals. Newhouse (1996) highlights how a mixed payment system, which uses both prospective and retrospective information, permits a tradeoff between selection and efficiency in production. Risk sharing is potentially attractive as a complementary strategy to risk adjustment. I focus on cases in which all of the risk sharing is between the regulator and the plan, and all of the risk adjusted payment is prospective from the regulator to the plan. Hence I focus on payment strategies that combine two regimes: prospective capitation payments (which may or may not be risk adjusted) and payments based on actual costs. Concurrent or retrospective risk adjustment is not analyzed in this paper.

A small but significant literature has examined how risk sharing can be used to reduce selection incentives while preserving incentives for efficiency.[2] Van Barneveld et al. (1996) introduce the concept of "Risk Sharing for High Risks" in which health plans are invited to identify at the beginning of each year a set of people enrolled in their plan that will be reinsured. The plans are in effect paid on a FFS basis for these reinsured individuals, but most accept capitation payments for the remainder of their enrollees. This innovation undermines the selection incentive problem by enabling plans to use their private information not to skimp, or distort services, or dump enrollees, but rather to use it to avoid the losses on those expected to be bad risks. In other work, they examine selection and efficiency incentives in a series of papers on which this paper builds.

In this paper, the regulator (principal), that receives information with lower quality than consumers about their level of severity, has to design a reimbursement scheme to health plans (agents). The aim is to provide a framework in which different payment systems (using risk adjustment and risk sharing strategies) can be compared in terms of social welfare. Risk adjustment strategies use only prospective information, while risk sharing uses *ex post* information based on actual costs, reducing the incentives for selection, but also for efficiency. The welfare loss is produced by the joint effect of both service level selection and inefficiency. This approach differs from the optimal risk adjustment discussed in Glazer and McGuire (2002a) in that the regulator does not impose efficiency in the schedule offered by HMOs and then minimize selection. Instead, in the theoretical model developed here, the regulator maximizes social welfare, or equivalently, minimizes the welfare loss produced by the inefficient schedule of services offered by two types of plans, HMO and FFS, taking into account the enrollment at each type of plan with selective behavior, and then, minimizes the profits that give incentives to select. The regulator reimburses retrospectively for all costs incurred to FFS plans, and reduces the welfare loss through the choice of the reimbursement system for HMO health plans, affecting the incentives for selection and efficiency in the HMO sector. Once the payment system is known, FFS and HMO health plans announce their schedule of health services offered to consumers. FFS are risk neutral and offer as many services as are demanded by consumers. HMO plans maximize profits by the selection of expected profitable consumers and choose the schedule of services that attract them for enrollment. Consumers differ not only in the level of severity of a chronic illness but also in the preferences for the different types of plans. They receive information on the schedule of services offered by the two types of plans, and

[2] Reviews are provided in Van de Ven and Ellis (2000), and Van Barneveld (2000).

depending on their severity and preferences, consumers decide whether to enroll in a FFS or HMO health plan. Neither risk sharing nor risk adjustment strategies are clearly dominant. However, the risk adjustment and risk sharing strategies that minimize the welfare loss are superior to a pure capitated system using no risk adjustment and to the conventional risk adjustment strategy.

As others have shown, selection efforts arise due to a key assymetry, such that the regulator is unable to remove the incentive for health plans to try to avoid certain unprofitable enrollees (whether due to poorer quality information, or regulations that prevent the regulator from using available information). In the particular model that I have developed, the HMO does not need to observe the same information that the consumer uses to make health plan choices, it only needs to know the structure of information that the consumer uses and the consumers utility function (or equivalently, the consumers demand curve). Thus, the HMO does not actually need to know whether consumer A has a signal to be high or low cost, only that people who have a high cost signal (known to them) will react in a certain way to the offering of a certain schedule of acute and chronic care. The HMO designs this schedule so that patients will sort themselves in a particular, utility maximizing way, which is also the firm's profit. Note that if instead I had modeled enrollee dumping rather than skimping, then the HMO would have to know the signals of individual enrollees. For simplicity, I assume that the HMO has the same information as the consumers, but in reality, the results would be the same if the HMO only knows the structure of decision making and the proportions of high and low cost consumers.

The rest of this paper after this introduction is organized as follows. Section 2 introduces the agents of the model and the information structure, that is, when they choose, and what they know at the time of taking their decision. Section 3 presents the three stage model in which regulator selects the payment system for plans, health plans choose the schedule of services offered, and consumers decide whether they prefer to enroll in a HMO or a FFS plan. Section 4 shows the correspondence between the theory model and the simulation model. Section 5 presents the data and simulation methods. Section 6 shows the simulation results, and finally, section 7 concludes.

2 Agents and information structure

2.1 Agents

There are three types of agents in the health care market: consumers, health plans, and regulator. Consumers differ in two dimensions: severity in a chronic illness, and the preferences on the rate of substitution between health services and money. Health plans are of two classes: health maintenance organization (HMO) and fee-for-service (FFS). There is only one benevolent regulator.

Consumers

The first source of heterogeneity in consumers comes from the illnesses they suffer. They can experience two different illnesses, acute and chronic. Acute illness a, is suffered by all consumers with the same severity. Thus, all consumers obtain the same utility from each dollar spent in the health services provided for this illness. Differently, the utility obtained from the care in chronic illness b, depends on their level of severity θ. Let us assume that there are two different levels of severity in the chronic illness, low or high, indexed respectively by L and H $(\theta^L < \theta^H)$ indicating the expected total spending in health services, which is higher for consumers with high severity and is given by $a + \theta b$ where $\theta = \left\{\theta^H, \theta^L\right\}$. A proportion η of the consumers suffer the chronic illness with low severity, and a proportion $1 - \eta$ with high severity. In the model presented here, I assume that consumers know their level of severity and so, they can calculate their expected cost of health services given a determined schedule of services offered. However, they do not know the realization of the actual or future cost that also will depend on unpredictable shocks.

The preferences between health services and money represent the second source of heterogeneity in consumers. I assume that the rate of substitution between the two goods λ follows a uniform distribution $\lambda \sim U[0, 1]$ independent on the level of severity so that consumers in the extremes with $\lambda = 0$ would only care about health services, and consumers with $\lambda = 1$ would equally weight the utility obtained from money and from health services.[3] Preferences are known by consumers

[3] This type of framework allow different interpretations for the preferences. Thus, they might represent the intensity on tastes for certain services or style of care. For instance, the lower λ is, the less important money is, and thus, there is more predisposition to enroll the FFS which means a higher importance given to the flexibility or independence from the HMO network. This preferences also can be understood in a Hotelling framework as a measure of the distance to each provider. Under this interpretation, the higher λ is, the closer the consumer is to

and enter in their utility function.

Health Plans

There are two types of health plans: HMO and FFS. I assume that in each geographical market, there is a large number of FFS, but I allow the existence of either only one monopolistic HMO or a number of HMO plans that collude and behave equally as a monopoly[4] or perfect competition. The only competition facing an HMO plan comes from the FFS sector. The representative HMO is a strategic profit maximizing agent that chooses a schedule of quality in the health services offered to consumers depending on their level of severity. Here, the cost is directly proportional to the quality chosen. The HMO is reimbursed by the regulator under a determined payment policy that can only use *ex ante* health signals or *ex post* actual costs. HMO health plans in this model do not need to know the level of severity of consumers. However, it is enough for health plans to know the structure of information that consumers use and their utility function when HMOs design the schedule for acute and chronic care using service level selection. Health plans do not know the realization of the preferences in consumers.

FFS health plans are reimbursed for all the cost of health services provided to consumers by the regulator. FFS plans also choose the schedule of quality in services they are willing to provide by offering as much as is demanded by consumers.

Regulator

The last agent in the health care market is the regulator. It is assumed that selection incentives stem from the fact that the regulator is unable to remove the incentive for health plans to try to avoid certain unprofitable enrollees. In this paper, there is perfect information for consumers regarding the level of severity in the chronic illness. It is not important the information that the regulator receives, what matters is the information that the regulator uses in the risk adjustment formulae.[5] I assume that the regulator is only allowed to use information contained in an imperfect signal on the level of severity observed by consumers (table 1). If the true severity level of the consumer is θ^H, the regulator receives the signal σ_1 with probability γ , and the signal σ_0 with

the HMO provider, and therefore, further from the FFS provider.

[4] The number of Medicare HMO health plans per county is very small. In Baker (1997), only 110 counties out of 3073 have Medicare HMO market penetration higher than 15%.

[5] Schokkaert and Van de Voorde (2004) differentiate between legitime and unlegitime risk adjusters.

probability $1-\gamma$. On the contrary, if the true level of severity is θ^L, the regulator receives the signal σ_1 with probability $1-\gamma$, and the signal σ_0 with probability γ.[6] The regulator does not have information on preferences.

Table 1: Probabilities of the imperfect signal used by the regulator

		Actual Severity	
		low	high
Signal	σ_0	γ	$1-\gamma$
	σ_1	$1-\gamma$	γ

2.2 Timing

All the decisions for period t are made at the end of period $t-1$ under the following scheme depicted in figure 1:

Stage 1: The regulator chooses the reimbursement payment system for HMO plans so as to maximize welfare or minimize a Social Welfare Loss produced by the inefficient schedule of services offered by the FFS and HMO plans and their enrollment rate. In order to do so, it can select one out of three possible payment systems: conventional risk adjustment (CRA) with its special case of No Risk Adjustment (NRA), minimum welfare loss risk adjustment (MWRA), and outlier risk sharing (ORS), taking into account the signals received from the market (σ_0 or σ_1) about the level of severity in the consumers or their actual cost. Through this choice, the regulator affects the incentives for selection and efficiency in the HMO, modifying the HMO schedule and reducing the welfare loss.

Stage 2: Health plans choose, given the payment method, the schedule of services offered to consumers. This schedule consists of a quality in the services offered for illnesses a and b, translated into amount of expected spending on each service, where the services for illness b can depend on the patient severity θ. The schedule, however does not depend on preferences λ, since they are unknown for health plans. Contract is therefore given by the schedule of expected spending in services a_H and θb_H in the case of the HMO health plans, and a_F and θb_F in the case of FFS

[6] The assumption that consumers have perfect information on the severity and the regulator does not, can be interpreted as if consumers and health plans can use information of better quality than the regulator.

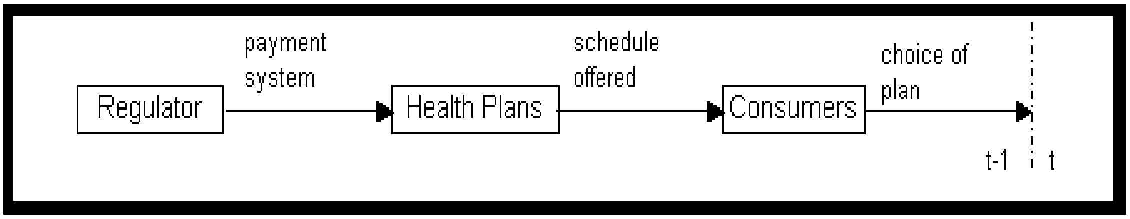

Figure 1: Timing in the decisions for period t are made at the end of period $t-1$

plans.

HMO plans choose the schedule $\{a_H, \theta b_H\}$ that maximizes expected profits given the reimbursement scheme and observing the structure of information used by consumers and their utility function.

At the same time, FFS plans have a dominant strategy, choosing the schedule of services $\{a_F, \theta b_F\}$ that maximizes the utility obtained by consumers. Thus, FFS choice can be thought as if FFS plans solve their problem in a degenerate stage instead of in the second stage in this model.

Stage 3: Finally, consumers maximize utility by choosing the plan in which they are willing to enroll (HMO or FFS) depending on their level of severity and preferences. I assume that there is an open enrollment period, so that they cannot be dumped by health plans. Their decision is taken by comparing the utility obtained from each health plan type, given both announced schedules.

This model is solved using backwards induction. First I solve the consumer utility maximization problem. Second, and given the choice of plan, I solve the selection of schedule of services by health plans. Lastly, the regulator minimizes the welfare loss with the constraints in the behavior from the other agents.

3 The Model

3.1 First stage: Consumers

Consumers' utility function depends on four components: their level of severity in the chronic illness ($\theta = \left\{\theta^L, \theta^H\right\}$), the schedule of services offered by the chosen plan ($\{a_F, b_F(\theta)\}$ and $\{a_H, b_H(\theta)\}$), the weight given to health services and money on the preferences λ, and the premiums paid to the regulator for the enrollment at each type of health plan, represented by $p = \{p_F, p_H\}$. I assume that the premium is zero if the consumer enrolls the HMO ($p_H = 0$). Differently, when a consumer enrolls in the FFS the premium is a proportion $\hat{s}$ of the expected cost, characterizing the existing demand side cost sharing differences between the FFS and HMO plans when making their choices. Thus, the utility function can be written as follows:

$$u_j(\theta^i, \lambda^i) = v_a(a_j) + v_b(\theta^i, b_j) - \lambda^i p_j(\theta^i) \tag{1}$$

where i is the index representing the consumer's level of severity ($\theta^i = \left\{\theta^L, \theta^H\right\}$) and preferences ($\lambda^i \in [0, 1]$), and j is the index for the health plan type ($j = F, H$).

Utility is given by a separable utility function in each type of service. Health services are normal *goods*. Therefore, *more* spending on health services for any of the illnesses is preferred to *less*. As a consequence, the first derivative of the utility function with respect to a_j and b_j is strictly positive. At the same time, I assume a diminishing marginal rate of substitution in spending in health services, meaning that given the acute illness a and a determined level of severity θ in the chronic illness b, initial units of consumption of each of the health service provide more utility than following units. I only consider consumption while the marginal utility of services is nonnegative. $v_a(a_j)$ represents the utility obtained from service a by plan j, and $v_b(\theta^i, b_j)$ the expected utility from service b by plan j for a consumer with a given level of severity defined by θ^i. The (expected) utility function for each of the health services is strictly concave on the amount of services ($v' > 0$ and $v'' < 0$) with the logarithmic specification $v_a(a_j) = \ln a_j$ and $v_b(\theta^i, b_j) = \theta^i \ln b_j$. Therefore, the utility function is characterized by:

$$u_j(\theta^i, \lambda^i) = \ln a_j + \theta^i \ln b_j - \lambda^i p_j(\theta^i) \tag{2}$$

Consumers, at the time of choosing the plan of enrollment, know their level of severity θ^i, preferences λ^i, and the schedule of spending offered by each health plan type. Hence, they maximize utility by comparing their utility levels at each plan type and selecting the highest. As a consequence, consumers with level of severity and preferences such that $u_H(\theta^i, \lambda^i) \geq u_F(\theta^i, \lambda^i)$ enroll in the HMO, and in the contrary case, they enroll in a FFS plan. Since the schedules are given at each health plan, this problems is equivalent to make the decision dependent on the level of preferences λ. Let $\tilde{\lambda}$ be the level of preferences that makes the consumer indifferent between two types of plans. Thus, if $\lambda^i < \tilde{\lambda}$ the consumer will enroll in the FFS plan, and otherwise, in the HMO.

3.2 Second stage: Health plans

In order to model health plans behavior, I follow the shadow price approach in Keeler et al. (1998), extended by Frank, Glazer and McGuire (2000), under which Managed Care health plans ration the spending in health services that each consumer receives. This approach, being equivalent to a model with real prices for each service, follows the health economics literature on optimal risk adjustment (Glazer and McGuire, 2002a) and is able to capture strategies of rationing care different than pure demand side cost sharing.[7]

Using this approach, the schedule of spending in services offered by health plan j $\{a_j, \theta b_j\}$ is such that the marginal utility obtained by a consumer from each service is equal to the shadow price of the service. In this model there is a shadow price fixed by each health plan j for the acute

[7] Even though there is in principal an equivalence between modeling plans as choosing shadow prices or choosing qualities or levels of expenditure for each service and for each consumer type, there is an important analytical convenience for using shadow prices. In the model used here there are only two services and two types of consumer health statuses. I could model the HMO as choosing four service levels, but instead model the plan as choosing only two shadow prices. As the previous literature has found, this results in a significant simplification. This simplification is even greater if I were to model more than two types of consumers, as I do below in my simulation analysis. Because it generalizes easily as multiple consumer types are added, I choose to assume that HMOs choose a shadow price for each type of service, rather than a rationing quality for each service offered to each consumer type. Note that the FFS does not get to choose the relevant shadow prices for each service: they are chosen by the regulator through the proportional cost share (s), and consumers, not the health plan, decide on the quality of services demanded.

illness (q_j^a) and other for the chronic illness (q_j^b).[8]

$$v_a'(a_j) = \frac{1}{a_j} = q_j^a \tag{3a}$$

$$v_b'(\theta, b_j) = \frac{\theta}{b_j} = q_j^b \tag{3b}$$

From here we can obtain the schedule of expected spending for each service depending on the shadow price chosen by each health plan type:

$$a_j = \frac{1}{q_j^a} \quad \text{and} \quad \theta^i b_j = \frac{\theta^i}{q_j^b} \tag{4}$$

3.2.1 Degenerate Stage: Fee-for-service plans

FFS plans choose their schedule of spending in services $\{a_F, \theta b_F\}$. As mentioned above, they are fully reimbursed retrospectively in a cost based formula for all costs incurred, including wages. Therefore, FFS plans do not have incentives to select and they accept any consumer willing to enroll. Furthermore, competing FFS provide as many services as demanded by consumers, whose demand function of health services is given by their utility function. Note that consumers obtain utility from spending on health services as long as the marginal utility function of the spending on each type of service is positive. Thus, the competitive solution for FFS plans is to offer an schedule in which the marginal utility equals the marginal cost paid by the consumers, which is itself given by the coinsurance rate $\hat{s}$ and the marginal rate of substitution between health services and income λ. As a consequence, the solution to the FFS maximization problem neither depends on the choice of the enrollment by consumers, nor the schedule offered by HMO plans, but only depends on the total retrospective reimbursement they receive from the regulator and the proportion of cost share for which FFS enrollees are responsible. Thus, the decision of the schedule of spending on services offered by FFS can be studied as if FFS choice is taken in a degenerate stage, and the solution is fixed and known by all agents. Analitically, the demand of spending on service a is given by:

$$\max_a \Psi^{FFS}(a, \hat{s}) = v(a) - \hat{s}\bar{\lambda} * a \tag{5a}$$

[8] Another approach to the topic of shadow prices is given by Ma (2003) in which health plans choose a different set of shadow prices for each group of consumers, by equalizing the marginal utility across services for each group.

where the first order condition is:

$$\frac{\partial \Psi^{FFS}(a, \hat{s})}{\partial a} = v'(a) - \hat{s}\bar{\lambda} = 0 \tag{5b}$$

From the consumers' perspective, the optimal schedule of spending on service a in FFS is such that $v'(a) = \hat{s}\bar{\lambda}$, where $\bar{\lambda}$ represents the average consumer marginal rate of substitution indicating the preferences on health services and income for the average consumer. Thus, the FFS sector chooses the schedule based on a real price of the health services, a demand side cost share, because there will not be any consumer demanding more spending. Hence, consumers are the agents rationing the schedule of services offered by FFS. Even though this is actually a demand side cost share, I assume that competing FFS offer this desired level of care. The same analysis is valid for the schedule of spending for the chronic illness.

Because s fixed and the same for both services a and b, and given the assumption of its distribution $\bar{\lambda} = \frac{1}{2}$, the resulting schedule $\{a_F, b_F(\theta)\}$ is determined by the shadow and real price for both types of illnesses ($q_F^a = q_F^b = s, \;$ with $s = \frac{1}{2}\hat{s}$), and following equation 4, is given by $a_F = \frac{1}{s}$ and $\theta^i b_F = \frac{\theta^i}{s}$.

3.2.2 Second stage: Health Maintenance Organization plan

As mentioned, the HMO is a strategic profit maximizing agent who provides services in the health care market. It uses information on the structure of information that consumers use and their utility function (or equivalently, consumers demand curve). It does not observe the individual preferences given by parameter λ but knows its distribution. Given the existence in this model of an open enrollment period, the HMO cannot dump any consumer willing to enroll. As FFS plans, it offers a schedule of spending in health services on each illness affecting the choice of plans by consumers. However, differently than FFS, the HMO (and not consumers) rations the services provided, and only consumers willing to enroll will be in the plan.

The HMO plan is reimbursed by the regulator differently than FFS, through a payment policy that might include both prospective and *ex post* information. Once it knows the payment system (from stage 1), the set of expected profitable enrollees is determined and the HMO plan maximizes profits by choosing the shadow price for each of the services, q_H^a for the acute illness, and q_H^b

for the chronic illness. Shadow prices for both services are not necessarily equal. As in Frank, Glazer and McGuire, the HMO will try to attract expected profitable enrollees (with lower level of severity) by offering more spending for the acute illness than FFS plans, but avoiding the rest of consumers with higher level of severity, by offering a lower spending for the chronic illness (service-level distortion).[9]

When shadow prices are chosen by HMOs, the spending in services a_H and b_H follow immediately from the demand functions and as shown in equation 4, they are given by:

$$a_H = \frac{1}{q_H^a} \quad \text{and} \quad \theta b_H = \frac{\theta}{q_H^b} \tag{6}$$

From now on, and in order to simplify notation, I eliminate the subscript of F for FFS or H for HMO in the shadow prices. Instead, the shadow price in the FFS sector is given directly by s, and the shadow price for the HMO is given by q^a in the case of service a, and q^b in the case of service b. In general it is expected to find that $q^a < s < q^b$, which is to say that the HMO will distort services by increasing the supply of acute treatment and reducing the supply of chronic care relative to the FFS.

Proposition 1 *If there is service level distortion, and thus, $a_H > a_F$, and $b_F > b_H$. Then, $\frac{\partial \tilde{\lambda}}{\partial \theta} > 0$ and thus, $\tilde{\lambda}(\theta^H) > \tilde{\lambda}(\theta^L)$, which means that the indifferent level of preferences given by the rate of substitution $\tilde{\lambda}$ will be higher for consumers of high severity than for consumers of low severity.*

Proof. *Set first the indifferent level of preferences $\tilde{\lambda}$ at each level of severity θ:*

$$\begin{aligned} u_H(\theta, \tilde{\lambda}) &= u_F(\theta, \tilde{\lambda}) \\ v_a(a_H) + v_b(\theta, b_H) &= v_a(a_F) + v_b(\theta, b_F) - \tilde{\lambda}\hat{s}\left[a_F + \theta b_F\right] \ \text{. Now,} \\ \tilde{\lambda}(\theta, \hat{s}, a_H, b_H) &= \frac{\left[v_b(\theta, b_F) - v_b(\theta, b_H)\right] - \left[v_a(a_H) - v_a(a_F)\right]}{\hat{s}\left[a_F + \theta b_F\right]} \end{aligned}$$

Substituting a_H, a_F, b_H, and b_F for their values and rearranging terms, we can rewrite the

[9] When the HMO chooses different shadow prices for both illnesses, with $q_H^a < q_H^b$, the HMO contract is neither First Best nor Second Best. Although it is not clear whether the overall budget for health care is too big or too low, by overproviding services for the acute illness spending on acute illness will be too high, and by underproviding services for the chronic illness, spending on this type of illness will be too low (Glazer and McGuire, 2002a).

indifferent $\tilde{\lambda}$ *as:*

$$\tilde{\lambda}(\theta, \hat{s}, q^a, q^b) = \frac{\theta \ln \frac{2q^b}{\hat{s}} - \ln \frac{\hat{s}}{2q^a}}{2(1+\theta)} \text{ or equivalently } \tilde{\lambda}(\theta, s, q^a, q^b) = \frac{\theta \ln \frac{q^b}{s} - \ln \frac{s}{q^a}}{2(1+\theta)}$$

Taking derivatives,

$$\frac{\partial \tilde{\lambda}(\theta, \hat{s}, q^a, q^b)}{\partial \theta} = \frac{2\left[\ln \frac{2q^b}{\hat{s}} + \ln \frac{\hat{s}}{2q^a}\right]}{4\left(1+\theta\right)^2} > 0$$

Since, with service level distortion, $a_H > a_F$*, and* $b_F > b_H$*, and thus,* $q^a < s$ *and* $q^b > s$*, with* $s = \frac{1}{2}\hat{s}$*.* ■

Proposition 2 *As the shadow prices chosen by the HMO increases, the level of preferences which makes the consumer indifferent between the two types of plans will also increase. In other words, as the HMO offers a lower spending in health services, a higher proportion of consumers will enroll in the FFS plan. Analitically,* $\frac{\partial \tilde{\lambda}}{\partial q^a} > 0$ *and* $\frac{\partial \tilde{\lambda}}{\partial q^b} > 0$*.*

Proof. $\tilde{\lambda}(\theta, \hat{s}, q^a, q^b) = \frac{\theta \ln \frac{2q^b}{\hat{s}} - \ln \frac{\hat{s}}{2q^a}}{2(1+\theta)}$

$\frac{\partial \tilde{\lambda}(\theta, \hat{s}, q^a, q^b)}{\partial q^a} = \frac{1}{2(1+\theta)q^a} > 0$

$\frac{\partial \tilde{\lambda}(\theta, \hat{s}, q^a, q^b)}{\partial q^b} = \frac{\theta}{2(1+\theta)q^b} > 0$ ■

Proposition 3 *As the coinsurance rate increases,* $\tilde{\lambda}$ *decreases. In other words, the higher is the proportion of costs for which FFS enrollees are responsible, the lower is the level of preferences that makes the consumer prefer the enrollment in the HMO. Analitically,* $\frac{\partial \tilde{\lambda}}{\partial \hat{s}} < 0$*.*

Proof. $\tilde{\lambda}(\theta, \hat{s}, q^a, q^b) = \frac{\theta \ln \frac{2q^b}{\hat{s}} - \ln \frac{\hat{s}}{2q^a}}{2(1+\theta)}$

$\frac{\partial \tilde{\lambda}(\theta, \hat{s}, q^a, q^b)}{\partial q^a} = -\frac{1}{2\hat{s}} < 0.$ ■

HMO profit maximization problem

The representative HMO plan receives information on the reimbursement system R and observes the level of severity of consumers. Then, it chooses the shadow prices for both health services, comparing the expected profits it obtained by attracting different consumers.

HMO knows that in order to attract consumers, it has to offer an schedule of services such that the utility they obtain from the HMO is at least as high as the utility obtained in the FFS plan.

Thus, in the profit maximization problem, the HMO has to take into account the proportion of consumers of each level of severity that will be willing to enroll given their preferences and the schedule offered by HMO and FFS. The expected profit maximization problem is formulated as the difference between the reimbursement system and the expected cost of the services provided:

$$\begin{aligned}\max_{q^a,q^b}\Pi &= \eta\left[1-F(\tilde{\lambda}(\theta^L,s,q^a,q^b))\right]\left[R-\frac{1}{q^a}-\theta^L\frac{1}{q^b}\right]+\\ &+(1-\eta)\left[1-F(\tilde{\lambda}(\theta^H,s,q^a,q^b))\right]\left[R-\frac{1}{q^a}-\theta^H\frac{1}{q^b}\right]\end{aligned} \tag{7}$$

In this problem it is already included the enrollment constraint, since only consumers with $\lambda^i > \tilde{\lambda}(\theta^i, s, q^a, q^b)$ will enroll in the HMO plan. Other constraints are the nonnegativity in the shadow prices ($q^a, q^b \geq 0$), and a participation constraint determined by the nonnegativity of expected profits ($\Pi \geq 0$). Recall that η represents the proportion of consumers suffering the chronic illness with low severity. The first part of the objective function refers to the expected profit derived to low severity enrollees, and the second refers to high severity enrollees. Competition can be added in the model allowing a large number of HMOs to participate in the market. When that is the case, FFS is not the only competition that HMO bears, and health plans have to provide more services. The effect of the existence of competitive HMOs is translated into an increase of the intensity in care for acute services: from the maximum profit allocation, some competitive HMOs will offer more intensity of care in acute services in order to attract the low severity consumers and still earn profits. The competitive equilibrium is achieved when there is no incentive to offer more intensity because that would provide expected losses for the associated enrollment, that is, when it holds the competitive constraint $\Pi\left(q^{a*} \pm \varepsilon, q^{b*} \pm \varepsilon\right) < 0$, being $q^{a*}$and $q^{b*}$the profit maximizing shadow prices (a higher schedule would provide negative profits and a lower schedule would attract no enrollees).

3.3 Third stage: Regulator

The regulator is the agent in the health care market choosing the reimbursement system R for the HMO. This choice is made in order to maximize social welfare, or equally minimize the social welfare loss, and reduce the incentives to select given by profits. Welfare loss is produced by the inefficient provision of spending on services in HMO and FFS. The regulator uses in this choice the imperfect signal on the level of severity described in table 1, and the cost structure.

Service efficiency

Social welfare is related to the principles of equity and efficiency. Equity can be understood as a distribution of the burden of health care costs equally among all individuals (Glazer and McGuire, 2000). In this model, I have assumed that FFS enrollees of any level of severity pay a proportion of cost to the regulator as reflected in the utility function, and no HMO enrollee pays a premium. Therefore, the cost share is the same for consumers with identical level of severity and enrolled in the same plan. As a consequence, there is equity in the health care market, and the regulator maximizes social welfare through the efficiency in the schedule of services.

The regulator maximizes an objective function which is increasing in the consumers' utility from services, but decreasing in the cost of those services. In the case of service a, the efficient provision of spending is given by:

$$\max_{a} \ v(a) - a \tag{8a}$$

where, as above, $v'(a)$ represents the marginal utility and demand function for spending in service a. For service b, an analogous objective function is needed for each level of severity.

$$\max_{b} \ v(\theta, b) - \theta b, \qquad \theta \in \left\{ \theta^L, \ \theta^H \right\} \tag{8b}$$

where $v'(\theta, b)$ represents marginal utility and demand function for spending in service b given the level of severity θ. Glazer and McGuire (2000) show that the efficient levels of spending in treatment for the two diseases are defined by shadow prices equal to one, where the marginal benefit obtained from every dollar spent in treatment equalizes its marginal cost of treatment for each service.

$$v_a'(a^e) = 1 \tag{9a}$$

$$v_b'(\theta, b^e) = 1, \ \theta \in \left\{ \theta^L, \ \theta^H \right\} \tag{9b}$$

Where a^e and θb^e represent the efficient schedule of spending in services and are calculated to be $a^e = 1$ and $\theta b^e = \theta$, with a different efficient expected spending for each level of severity.

Payment systems

The three reimbursement systems considered in this paper are the following:

- *Conventional risk adjustment (CRA)*: the regulator is only allowed to reimburse the HMO by using the information contained in the signal in a regression-based approach. The payment schedule is based on the expected cost for the FFS sector and it is the same for all consumers belonging to the same risk group given the information to each group.[10] There are two different reimbursement amounts for the two risk groups of consumers depending on the *ex ante* signals obtained by the regulator (σ_0 and σ_1).

$$R^{CRA}_{\sigma_0} = a_F + \left[(1-\gamma)\theta^H + \gamma\theta^L\right] b_F = \mu_{\sigma_0} \tag{10}$$
$$R^{CRA}_{\sigma_1} = a_F + \left[\gamma\theta^H + (1-\gamma)\theta^L\right] b_F = \mu_{\sigma_1}$$

Note that a special case occurs when the signal is completely uninformative ($\gamma = \frac{1}{2}$), or simply when the regulator does not use any information. Then the payment system applied is No Risk Adjustment (NRA) and the reimbursement corresponding to all consumers is the grand mean of costs ($R^{NRA} = \mu$). If the regulator could identify k risk groups, there would be a R^{CRA}_k for each group equivalent to the average expected cost of that group μ_k.

- *Minimum Welfare loss Risk Adjustment (MWRA)*: the regulator reimburses the HMO using the *ex ante* signals, but instead of using regression-based risk adjusters, it is allowed to use biased estimators with the weights ϕ_a and ϕ_b that minimize a welfare loss function. The solution is optimal within this framework. However, it is useful to clarify the difference with the optimal risk adjustment formula in Glazer and McGuire (2002a). While the ORA payment assures efficiency in the HMO and then minimizes selection, the MWRA payment in this framework does not, and uses a different approach by minimizing the total welfare loss produced by both inefficiency and selection given the enrollment in the HMO and the FFS sector:

$$R^{MWRA}_{\sigma_0} = \phi_a a_F + \phi_b \left[(1-\gamma)\theta^H + \gamma\theta^L\right] b_F \tag{11}$$
$$R^{MWRA}_{\sigma_1} = \phi_a a_F + \phi_b \left[\gamma\theta^H + (1-\gamma)\theta^L\right] b_F$$

When the regulator can only divide consumers in k risk groups with respect to total need

[10] This case reflects the Medicare case, where the regulator pays each health plan the average spending.

of care, this is equivalent to use a parameter of distortion δ in the payment with respect to the unbiased conventional risk adjustment payment (information set A). Then, the reimbursement is given by

$R_k^{MWRAA} = R_k^{CRAA} + \delta \left(R_k^{CRAA} - R^{NRA} \right) = \mu_k + \delta \left(\mu_k - \mu \right)$. However, if the regulator used information separately on acute and chronic illness, then she can distort the payment with a biased formula through δ_1 and δ_2 as follows. In the simulations we will name this information set B: $R_k^{MWRAB} = (1 - \delta_1) a_k + (1 + \delta_2) b_k$.

It is important to note that given the role of the distortion parameters, the MWRA formula is a biased payment system with respect to the regression based approach.

- ***Outlier Risk Sharing (ORS)*:** the regulator uses *ex post* information in order to retrospectively reimburse the HMO for some of the cost incurred (R_r) in high cost enrollees, and also there is a part of the reimbursement that is prospective (R_p). The prospective payment is based on a fraction (ω^p) of the no risk adjustment payment. The retrospective payment is used to compensate the HMO for high severity consumers and consists of a proportion (ω^r) of the difference between actual cost and a threshold T determined by the regulator.

$$R_p^{ORS} = \omega^p R^{NRA} \quad \text{and} \quad R_r^{ORS} = \omega^r (C^i - T) \quad \text{iff} \;\; C^i \geq T \tag{12}$$

where $R^{NRA} = \mu$ (grand mean of the cost) and C^i represents the actual cost of consumer i.

Welfare loss minimization problem

After describing the choice of spending in services by health plans, the efficient provision of services, and the payment systems it is possible to characterize the regulator's welfare loss minimization problem. First, the analytical formulae for the different schedules of services offered by plans and the efficient contract desired by the regulator are compared. FFS health plans are not efficient and have a shadow price ($0 < s < 1$) set by the coinsurance level. As a consequence, they overprovide spending (quality) in services ($a_F > a^e$ and $b_F > b^e$ for both levels of severity). Being $a^e = 1$ and $a_F = \frac{1}{s} a^e$, and $\theta b^e = \theta$ and $\theta b_F = \frac{1}{s} \theta b^e$(with $s < 1$) the proportion of overprovision in the FFS is $\frac{1-s}{s}$ in the case of both acute and chronic illnesses. Let us rename $q^a = \frac{1}{\alpha} s$ and $q^b = \frac{1}{\beta} s$; now we can compare the schedules provided by HMO and FFS through the parameters of over or under provision α and β, so that $a_H = \alpha a_F$ and $b_H = \beta b_F$, and the welfare loss in the HMO through the comparison of the efficient and HMO schedule. Therefore, the rationing parameters q^a and q^b can be converted in α and β, and for tractability, are assumed to be the same for all acute and chronic diseases respectively.[11]

$$a_F = \left(1 + \frac{1-s}{s}\right) a^e; \qquad b_F = \left(1 + \frac{1-s}{s}\right) b^e \tag{13a}$$

$$a_H = \alpha \left(1 + \frac{1-s}{s}\right) a^e; \qquad b_H = \beta \left(1 + \frac{1-s}{s}\right) b^e \tag{13b}$$

If there is selection, it is generally expected that $\alpha > 1$ and $\beta < 1$ ($q^a < s < q^b$) which is equivalent to the service-level selection assumption in Cao and McGuire (2003): with a shadow price holding $q^a < s < 1$, FFS allocation (a_F) is closer to the efficient provision (a^e) and welfare loss is lower than in the case of the HMO allocation (a_H). The contrary happens for the chronic illness. Figure 2 presents the offer of spending by health plans and the efficient allocation in the case of the acute illness with service-level distortion, and displays the welfare loss. Figure 3 does the same exercise for spending in the chronic service for consumers with both levels of severity. For each consumer enrolling a FFS health plan, the welfare loss incurred in the services offered is measured by the vertically striped area between the amount of services offered by health plans and the efficient welfare maximizing contract (given by $\int_1^{\frac{1}{s}} \left(1 - \frac{1}{x}\right) dx$ for service a, and by $\int_\theta^{\frac{\theta}{s}} \left(1 - \frac{\theta}{x}\right) dx$ for service b) and the welfare loss incurred by consumers enrolling an HMO is given by the horizontally striped area (given by $\int_1^{\frac{1}{q^a}} \left(1 - \frac{1}{x}\right) dx$ for service a and by $\int_\theta^{\frac{\theta}{q^a}} \left(1 - \frac{\theta}{x}\right) dx$ for service b). As it can be seen in both figures, when the HMO over (under) provides services with respect to the FFS schedule, the welfare loss is greater (lower). Although the figures reflect the case with service-level distortion, the solution of this model shows under which payment systems this assumption holds.

The regulator modifies HMO's behavior by choosing the payment system that minimizes the social welfare loss due to deviations from the contracts offered by providers and the efficient allo-

[11] Although Cao and McGuire (2003) does not indentify an "average" service level distortion in their data, the assumption of having one average service distortion for acute illnesses and other for chronic illnesses simplifies the behavior in the model and simulations in this paper. It is not clear that adding more complexity would result in further insights.

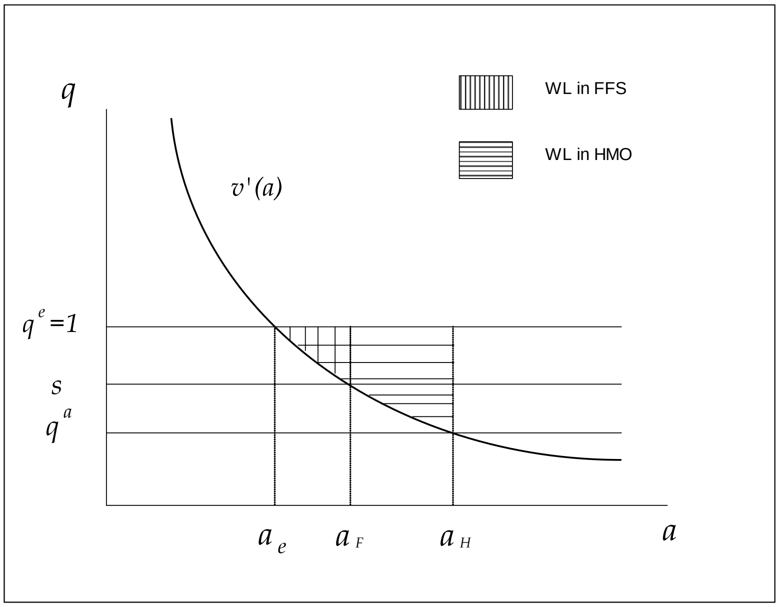

Figure 2: Welfare loss in service a from HMO and FFS

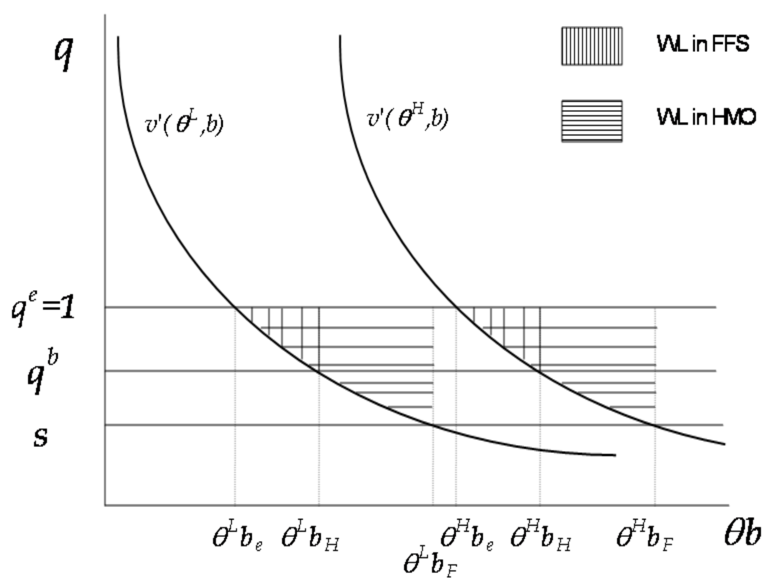

Figure 3: Figure 3: Welfare loss in service b from HMO and FFS for each level of severity

cation. Consumers maximize utility; however they cannot choose the efficient treatment, but only the contracts offered by HMOs and the FFS plans $\{a_H, b_H\}$ and $\{a_F, b_F\}$.

The Welfare Loss function is calculated as the sum of the welfare loss obtained from each type of health service, acute and chronic, ($WL = WL^a + WL^b$), taking into account the enrollment constraints, as follows:

$$\begin{aligned} WL^a &= \left[E_F^H + E_F^L\right]\left[(\frac{1}{s}-1) - \ln\frac{1}{s}\right] + \left[E_H^H + E_H^L\right]\left[(\frac{1}{q^a}-1) - \ln\frac{1}{q^a}\right] \quad (14a) \\ WL^b &= \theta^H E_F^H\left[(\frac{1}{s}-1) - \ln\frac{1}{s}\right] + \theta^L E_F^L\left[(\frac{1}{s}-1) - \ln\frac{1}{s}\right] + \\ &\quad +\theta^H E_H^H\left[(\frac{1}{q^b}-1) - \ln\frac{1}{q^b}\right] + \theta^L E_H^L\left[(\frac{1}{q^b}-1) - \ln\frac{1}{q^b}\right] \quad (14b) \end{aligned}$$

being $E_H^i = (1-\eta)(1-F(\tilde{\lambda}(\theta^i, s, q^a, q^b)))$ the proportion of consumers enrolling the HMO with severity $i = \{H, L\}$ and $E_F^i = (1-\eta)F(\tilde{\lambda}(\theta^i, s, q^a, q^b))$ the proportion of consumers enrolling the FFS with severity $i = \{H, L\}$, and therefore $E_j^H + E_j^L$ the enrollment rate in health plan j. Rearranging terms by health provider, WL_j represents the welfare loss given the enrollment rate in health plan j:

$$\begin{aligned} WL_F &= \left[(1+\theta^H)E_F^H + (1+\theta^L)E_F^L\right]\left[(\frac{1}{s}-1) - \ln\frac{1}{s}\right] \quad (15a) \\ WL_H &= \left[E_H^H + E_H^L\right]\left[(\frac{1}{q^a}-1) - \ln\frac{1}{q^a}\right] + \left[\theta^H E_H^H + \theta^L E_H^L\right]\left[(\frac{1}{q^b}-1) - \ln\frac{1}{q^b}\right] \quad (15b) \end{aligned}$$

The regulator compares the result of the welfare loss from the different payment systems taking as a constraint the maximization utility problem in the choice of plan of enrollment by consumers (included in the indifferent level of preferences $\tilde{\lambda}$ and the profit maximization problem by HMO health plans). The regulator anticipates the behavior from the HMOs profit maximization problem and consumer choice under each payment system, and designs the payment system so as to minimize the welfare loss.

This model allows to evaluate different objective functions for the regulator, and not only the minimization of the total welfare loss. It is also possible to minimize the service distortion or the welfare loss by HMO enrollee and thus the minimization of the cross subsidarization process between HMO and FFS per different types of individuals. When that is the goal, the regulator needs to calculate the welfare loss per enrollee in the HMO sector.[12]

[12] As a special case, we can assume that FFS are efficient, with no overprovision of spending. Then, by minimizing

4 Data and Simulation Methods

This paper utilizes data from the 5% Standard Analytical file for Medicare beneficiaries. This data contains information on 1,417,005 FFS enrollees in Medicare during years 1996 and 1997.

Demographic information is composed by gender and age (on January 1, 1997) on 22 age/gender cells which compose the risk groups described in the model section and is used by the regulator. There is also data on the annualized total covered expenses in 1997, and the event of suffering six chronic illnesses during year 1996. The chronic illnesses are cancer (Metastatic Cancer and Acute Leukemia), diabetes, Congestive Heart Failure (CHF), stroke (Cerebral Hemorrhage, Ischemic or Unspecified Stroke), COPD (Cystic Fibrosis or Chronic Obstructive Pulmonary Disease), and renal failure (Kidney Transplant Status, End Stage Renal Disease, Dialysis Status, or Renal Failure). Those six chronic conditions are the multiple services that are potentially distorted by the HMO sector in the model developed here.

Table 2 presents the descriptive statistics of the variables used. The average spending (grand mean), used by the regulator for the NRA payment system (as a especial case of CRA) is of $6,944. There is a 41.77% of males in the sample, and most of consumers in the sample are in the age between 65 and 84 years old (76.76%) corresponding to FFS Medicare beneficiaries. Among the chronic diseases, diabetes is the most suffered by consumers, followed by COPD, and CHF. A lower number of enrollees suffer stroke, renal disease and cancer.

In this model it is key how much information the regulator has in cotrast to consumers and health plans. As mentioned in the description of the payment systems, there are two different information sets that can be compared. Information set A corresponds to the case in which the regulator can divide population in k groups, but cannot observe information independently for acute or chronic illnesses. As a consequence there is only one δ parameter in the formula. Differently, Information set B corresponds to the case in which the regulator can observe signals for acute and chronic conditions per individual and therefore may use two distortion parameters δ_1 and δ_2 in the formula.

(and eliminating) service distortion, the welfare loss would be minimized and the payment formula would equal the Optimal Risk Adjustment formula in the literature.

In the simulations, I am going to assume different values for γ, the parameter providing the difference between the quality of the information observed by the regulator related to consumers and health plans. Tables 3 and 4 present the estimation of the spending in health care by the regulator and health plans. Table 3 refers to the case in which the regulator uses only demographic information, that is, no information on the chronic condition. This is equivalent in the model to $\gamma = 0.5$, and therefore, there is only estimation of costs for all consumers in each demographic cell. In other words, the regultator assigns the same probability of suffering a chronic illness to consumers providing signal σ_0 and σ_1 and there are 22 CRA payments. Thus, when $\gamma > 0.5$ there are 44 CRA payments, two per each of the 22 demographic cells. Table 4 presents the estimation of cost for health plans, when $\gamma = 1$. The CRA payment is based on the estimate of the expected spending in the FFS sector (without service distortion) using different values for γ, because all the risk adjustment payments are based on FFS expected costs. If the regulator used this expectations, there would not be incentives for selection. However, with $\gamma < 1$, health plans and consumers manage information with higher quality and in this model may take advantage of it through the service level selection (skimping) strategy. Table 5 presents how the R^2 of the predictive model increases with the quality of the information from $R^2 = 0.0103$ when it only uses demographic information to $R^2 = 0.0345$ when $\gamma = 0.95$, with the limit of $R^2 = 0.0413$ that correspond to the predictive model for health plans and consumers ($\gamma = 1$).

At the same time the analysis compare the effect of different coinsurance rates s providing the overprovision in the FFS sector, in order to test for higher and lower differences in premiums and welfare loss in different sectors.

When using NRA or CRA payment systems, the regulator cannot take a strategic behavior (there is no parameter she can choose) and only can estimate the welfare loss resulting from the HMOs choice of intensity in care and their expected profits. However, when the regulator uses minimum welfare loss risk adjustment (MWRA), the payment can be distorted with respect to the unbiased estimator given in CRA through the distortion parameter δ (information set A). The simulations analyzes whether the choice of intensity by HMOs changes with δ. In the case in which the regulator only uses demographic information, $|\delta| \in \{10, 20, 30, 40, 50\}$. If the regulator uses information set B, there are two parameters of distortion in the payment, one for spending in acute care (δ_1) and other for spending in chronic care (δ_2). I simulate the results for the different combinations of distortion in the payment relative to conventional risk adjustment when $\delta_1, \delta_2 \in \{5, 10, 15, 20\}$ and results are presented in table 6.

The last payment system analyzed in this study is outlier risk sharing (ORS), composed by a prospective payment, and a proportion of the actual cost beyond a threshold. There are three parameters that the regulator can choose: the prospective payment, the threshold, and the proportion of retrospective reimbursement. With respect to the first one, in the simulations I use four different lump sum amounts for the prospective payment (PP), which correspond to $\omega^p = \{80\%, 85\%, 90\%, 95\%\}$ of the no risk adjustment payment (grand mean of costs), and four different proportions $\omega^r = \{5\%, 10\%, 15\%, 20\%\}$ of the cost beyond the threshold reimbursed by the regulator. Regarding the threshold T, determining whether a consumer becomes an outlier or not, I select the cost corresponding to the 75, 80, and 85 percentiles of the cost ($T_{75} = 4924$, $T_{80} = 7271$, and $T_{85} = 11519$). A threshold lower than the corresponding to the 75 percentile would derive in the case in which consumers belonging to almost all demographic cells and levels of severity would be considered outliers. In order to obtain the probability of being an outlier given the level of severity, I use the assumption that the cost at each level of severity follows a normal distribution with mean the prediction obtained with the information of the consumers and its standard deviation (σ). Thus,

$$\Pr(C_H > T \mid \theta) = 1 - \Phi\left(\frac{T - \hat{C}_H}{\sigma}\right) \tag{16}$$

Once cost and reimbursement are characterized, HMO health plans can formulate the profit maximization problem in which they choose the service level distortion through parameters α and β, subject to the constrains of enrollment from consumers given by their utility function, and participation. As mentioned, the analysis presented here includes the case of perfect competition, which includes the competitive constraint translated into $\Pi(\alpha^* + \varepsilon, \beta^* + \varepsilon) < 0$. Once α and β are obtained, HMO profits, enrollment rates and welfare loss can be calculated.

It is important to examine the different combinations of values for the distortion parameters in which the HMO will maximize profits given that the reimbursement schemes based on costs in the FFS sector. First, any combination containing $\alpha \geq 1$ and $\beta > 1$, or $\alpha > 1$ and $\beta \geq 1$ (overprovision of spending with respect to FFS in at least one of the types of care and at least the same spending in the other), the HMO would obtain the enrollment of all consumers but negative profits. Second, if $\alpha \leq 1$ and $\beta < 1$, or $\alpha < 1$ and $\beta \leq 1$ (underprovision of spending) even if that would be a profitable schedule for the HMO, no consumer will be willing to enroll. Third, the special case in which $\alpha = \beta = 1$ occurs when the schedule offered by the HMO is identical to the one offered by the FFS sector, and thus, there are zero expected profits. Fourth, $\alpha < 1$ and $\beta > 1$ attracts to a proportion of consumers, but given the utility function, those are the high cost consumers, and thus, this combination of service distortion also produces expected losses. Finally, the only combination of service distortion that can provide positive expected profits and enrollment occurs if the HMO is able to attract consumers by overproviding spending in services for the acute care and underproviding spending in services for the chronic care, with $\alpha > 1$ and $\beta < 1$.

Result 2: In this model there are incentives for the service level distortion type of selection.

5 Simulation results

No risk adjustment is the payment system in which the regulator reimburses the same amount of money to the HMO health plans for any enrollee no matter their age, gender, or incidence of chronic illness. The payment consists of the grand mean of all costs, and thus, the regulator does not use any information or risk adjusters. If there was a unique HMO maximizing profits, the schedule of services offered would be determined by $\alpha = 1.001$ and $\beta_{NRA} = 0.992$, and there would be an enrollment 970,123 enrollees (all females with no chronic illness, and also males with no chronic illness and being less than 85 years old) with a profit of $2,252 million. However, allowing competition in the market, the intensity of acute services offered is increased until $\alpha_{NRA} = 1.01$. The new enrollment is of 1,326,682 (93.63% of all the population, including consumers with one and more than one chronic conditions at the end of year one) and the profits are reduced to $211 million. The welfare loss resulting from that schedule and enrollment, is increasing in the proportion p of overprovision in the FFS sector, with a level in millions of $WL_{NRA} = 529.93$ when $p = 5\%$.

When the regulator uses CRA, she uses the information observed in a regression based approach, with unbiased estimators. Conventional risk adjustment reduces the service level distortion, when the regulator uses only demographic information, the new schedule of intensity in services is given by $\alpha_{CRAA} = 1.009$ and $\beta_{CRAA} = 0.993$ once competition is added in the model. The profit obtained by the HMOs is of $\Pi_{CRAA} = \$248$ million, which is greater than with NRA. However, the enrollment has increased to 1,355,768 consumers (95.68%), and the welfare loss has also decreased relative to NRA, $WL_{CRAA} = 526.4$. When the regulator uses the information set B, the quality of the information increases and the payment CRAB is better adjusted to the cost expected by health plans. As a result, the service selection is reduced relative to NRA and also relative to conventional risk adjustment with only demographic information (CRAA). The schedule of spending in services is provided by $\alpha_{CRAB} = 1.007$ and $\beta_{CRAB} = 0.995$. The enrollment is also increased relative to NRA and CRAA with 1,401,195 enrollees: all consumers but those 65-69 years old females and males with more than one chronic condition at the end of year 1. The ratio of enrollment is therefore of 98.88%. CRAB also reduces profits to $\Pi_{CRAB} = \$74$ million, and the welfare loss to $WL_{CRAB} = 520.6$.

The Minimum Welfare Loss risk adjustment formulae (MWRA) distort the unbiased estimators used in CRA using the parameter δ (information set A) and δ_1 and δ_2 (information set B). When the regulator only uses demographic information (MWRAA), for all the analyzed values of δ ($|\delta| \in \{10, 20, 30, 40, 50\}$) the service distortion is reduced at the same level than with CRA ($\alpha_{MWRAA} = 1.009$ and $\beta_{MWRAA} = 0.993$), and therefore, the same enrollment and welfare loss are obtained. However, through δ, the regulator can reduce the profits expected by HMOs while they still offer the same schedule. Although profits are higher with MWRAA than with NRA for the value of the parameter δ used in the simulations, minimum welfare loss risk adjustment is superior to both NRA and CRAA because the welfare loss is lower than with NRA, and being the same than with CRAA, profits are reduced. At the same time, $\delta < 0$ would make MWRAA profits to be lower than NRA. It is important to note that service selection has been reduced and more enrollees choose the HMO than with NRA. Thus the profit per HMO enrollee has decreased under MWRAA relative to NRA and CRAA.

When the regulator uses information set B with some signals on the incidence of chronic illnesses, the MWRAB formula is distorted with two different parameters. Table 7 examines the profits, enrollment and welfare results for different combinations of the distortion parameters δ_1 and δ_2 in the payment formula. As it can be seen, the lowest welfare loss with MWRAB ($WL_{MWRAB} = 5.16$) is lower than that obtained with MWRAA and the choice of HMOs supposes a lower service distortion ($\alpha_{MWRAB} = 1.006$ and $\beta_{MWRAB} = 0.995$). However, the ratio of enrollment for that choice of intensity in services is lower (91.15%). From those combinations of distortion in the payment producing the same result of enrollment, welfare and service distortion, the regulator would choose the one minimizing profits. As a consequence, $\Pi_{MWRAB} = \$14$ million with 15% of underpayment for the expected spending in acute services, and 10% of overpayment for the expected spending in chronic services, much lower than obtained with conventional risk adjustment. Some combinations of distortion in the payment have been omitted because they produce expected losses and no HMO would enter in the health care market. It is worth to note that if the regulator tried to minimize HMO expected profits, the combination of distortion in the payment formula would be of $\delta_1 = 5\%$ and $\delta_2 = 10\%$. As a consequence, service distortion and welfare would be slightly higher with $\alpha_{MWRAB} = 1.008$ (still lower than with NRA) and all consumers would enroll in the HMO.

Hence, minimum welfare loss risk adjustment is superior to no risk adjustment and conventional risk adjustment both relative to how they reduce the welfare loss and profits, and the incentives to select through service level distortion (table 7).

The last payment system simulated in this paper is outlier risk sharing (results are presented in table 8). The intensity of services offered and enrollment is increasing on the prospective payment and in the quality of the signal observed by the regulator (γ) but is decreasing in the threshold.

When the prospective payment (PP) consists of a 80% of the NRA reimbursement, and $\gamma = 5\%$ or 10%, no matter the threshold used, only females with age lower than 80 years old and males with age lower than 74 years with no chronic condition would enroll in the HMO. Thus, the number of enrollees is of 709,955 and the ratio of enrollment is of 50,10%. The resulting service distortion consists of a very low overprovision of intensity in acute services ($\alpha_{ORS} = 1.001$), lower than the obtained with risk adjustment strategies, but a higher underprovision of intensity in chronic services ($\beta_{ORS} = 0.988$). The welfare loss when the overprovision in the FFS sector is of 5%, $WL_{ORS} = 4.95$ which is lower than that obtained under risk adjustment strategies including minimum welfare loss risk adjustment, and the expected profit is of $\Pi_{ORS} = \$0.02$ million. Other solutions are feasible when $\gamma = 15\%$ or 20%, or when the prospective payment is a higher proportion of the NRA payment, offering a higher intensity in the acute service ($\alpha'_{ORS} = 1.01$) and obtaining a higher enrollment rate. However, the welfare loss increases with α, and if the regulator seeks to minimize the welfare loss, she will pay prospectively 80% of the grand mean, and there will be only a small proportion of retrospective payment in expectation, since almost all enrollees will have an expected cost lower than the threshold.

Lastly, I have also simulated the effect of a change in the proportion of overprovision in the FFS sector. Results are presented in tables 9 and 10 for the cases in which $p = 0\%, 5\%, 10\%, 15\%$, and 20%. The first case supposes that the FFS schedule is efficient. Total welfare loss is increasing on the proportion of overprovision p (inefficiency) in the FFS sector, and is always lower under outlier risk sharing than under risk adjustment. From all the risk adjustment strategies, minimum welfare loss risk adjustment reduces better the welfare loss when the regulator uses her high quality information set. Thus, in terms of total welfare loss, ORS results a better strategy than risk adjustment. This is explained by the higher overprovision of acute services and the higher rate of enrollment in the HMO, since all consumers need acute services.

However, it is important to note that given the difference in the number of enrollees, the welfare loss per HMO enrollee is lower under biased formula of minimum welfare loss risk adjustment than under outlier risk sharing. This measure is important in terms of equity in the provision of healthcare to all individuals. Table 11 shows how the choice of the regulator trying to minimize the welfare loss per enrollee in the HMO depends on the quality of the information she uses ($WL^i_{MWRAB} < WL^i_{ORS} < WL^i_{MWRAA}$) for any level of overprovision in the FFS sector. When the only information used by the regulator is demographic, the service distortion is lower under outlier risk sharing, in the sense that the difference in the overprovision of acute spending (MWRAA versus ORS) is higher than the difference in the underprovision of chronic spending taking into account the enrollment derived from each payment system. As a consequence, the welfare loss per enrollee in the HMO is lower under risk sharing than under risk adjustment. However, when the regulator uses better information, the service distortion is lower under risk adjustment, because the difference in the overprovision of acute spending (MWRAB versus ORS) is lower than the difference in the underprovision of chronic spending, and then, the welfare loss per enrollee in the HMO is lower under risk adjustment. As a conclusion, outlier risk sharing reduces better the total welfare loss given the inefficient schedule of services offered by FFS and HMOs. Differently, minimum welfare loss risk adjustment formula can reduce better the service level distortion and the welfare loss per enrollee in the HMO, promoting equity in the healthcare provision, and obtains a higher enrollment rate. This relationship is stronger with a higher quality in the information used by the regulator to configure the risk adjustment formula.

6 Conclusion

This paper has developed a model which is able to compare through simulation the effect in terms of social welfare of different payment policies on health maintenance organizations (HMOs) in the Medicare market, including both risk sharing and risk adjustment strategies (no risk adjustment, conventional risk adjustment, and minimum welfare loss risk adjustment). The primary focus of the simulations has been on the extent to which selection incentives are reduced as the signal used by the regulator becomes more informative, and on the conditions under which risk adjustment strategies are superior or inferior to risk sharing strategies.

The comparison of the quality of the information that matters in this model is between the information used by consumers and regulator. Health plans do not need to use any information on the health status of consumers, what they need to know is the structure of information used by consumers and how they choose. HMO health plans choose the intensities of services for both types of care provided that makes maximize profits given the consumers' problem and the competition in the HMO sector.

The simulations confirm that the minimum welfare loss risk adjustment formula is superior to both conventional and no risk adjustment. This difference increases with the quality of the information used by the regulator. When comparing outlier risk sharing with minimum welfare loss risk adjustment it is important to define the objective function used by the regulator: outlier risk sharing reduces better the total welfare loss resulting from the inefficient schedule of services offered given the enrollment in HMOs and FFS for that schedule. However, when the regulator seeks to minimize selection in the HMOs and thus, promote equity in the provision of healthcare, the individual welfare loss per enrollee has to be compared under the different payment systems. Here, the quality of the information used by the regulator plays a key role: outlier risk sharing reduces better the welfare loss per enrollee in the HMO and service level selection than minimum welfare loss risk adjustment when the quality of the information used by the regulator is low. However, minimum welfare loss risk adjustment behaves better than outlier risk sharing when the difference between the quality of the information used by consumers and regulator is lower.

The potential of both prospective risk adjustment and risk sharing strategies has to be understood in terms of the difference in the quality of information available for health plans and regulator. Although several models are being explored, more research needs also to be developed considering the use of hybrid risk adjustment in which both prospective risk adjustment and risk sharing strategies are used together under different objective functions.

References

[1] Ash, A., R.P. Ellis, G.C. Pope, et al. (2000): "Using Diagnoses to Describe Populations and Predict Costs." Health Care Financing Review 21(3): 7-28.

[2] Beebe, J.C. (1992): "An Outlier Pool for Medicare HMO Payments." Health Care Financing Review 14(1): 59-63.

[3] Brown, R.S., J.W. Bergeron, D.G. Clement, J.W. Hill and S. Retchin (1993): "Does Managed Care Work for Medicare? An Evaluation of the Medicare Risk Program for HMOs." Report under HCFA Contract Number 500-88-0006. Princeton, NJ. Mathematica Policy Research, Inc.

[4] Cao, Z. and T.G. McGuire (2002): "Service Level Selection by HMOs in Medicare." Working paper.

[5] Cao, Z. (2002): "Service level Self-Selection in Medicare HMO Enrollment." Working paper.

[6] Chapman, J.D. (1997): "Biased Enrollment and Risk Adjustment for Health Plans." Unpublished doctoral dissertation, Harvard University.

[7] Ellis, R.P. (1998): "Creaming, Skimping, and Dumping: Provider Competition on the Intensive and Extensive Margins." Journal of Health Economics 17(5): 537-555.

[8] Ellis, P. R., G. Pope, and L. Iezzoni, et al. (1996): "Diagnosis–Based Risk Adjustment for Medicare Capitation Payments." Health Care Financing Review 17(3): 101-128.

[9] Ellis, R.P. and T.G. McGuire (1986): "Provider Behavior under Prospective Reimbursement." Journal of Health Economics 5: 129-141.

[10] Ellis, Randall P., and T. G. McGuire (1990): "Optimal Payment Systems for Health Services." Journal of Health Economics 9(4): 375-396.

[11] Ellis, R.P. and T.G. McGuire (1993): "Supply-side and Demand-side Cost Sharing in Health Care." Journal of Economic Perspective 7: 135-151.

[12] Frank, R.G., J. Glazer, and T.G. McGuire (1998): "Measuring Adverse Selection in Managed Health Care." Journal of Health Economics 19: 829-854.

[13] García-Goñi, M., P. Ibern (2008): "Predictability of Drug Expenditures: An Application Using Morbidity Data". Health Economics 17: 119-126.

[14] Glazer, J., and T.G. McGuire (2000): "Optimal Risk Adjustment in Markets with Adverse Selection: An Application to Managed Care." The American Economic Review 90(4): 1055-1071.

[15] Glazer, J., and T.G. McGuire (2002a): "Setting Health Plan Premiums to Ensure Efficient Quality in Health Care: Minimum Variance Optimal Risk Adjustment." Journal of Public Economics 84: 153-173.

[16] Keeler, E.B., Carter, Trude (1988): "Insurance Aspects of DRG Outlier Payments." Journal of Health Economics 7: 193-214.

[17] Kravitz, R. L., S. Greenfueld, W. Rogers, et al. (1992): "Differences in the Mix of Patients among Medical Specialties and Systems of Care: Results from the Medical Outcomes Study." Journal of the American Medical Association 267: 1617-1623.

[18] Lichtenstein, R., J. W. Thomas, J. Adams-Watson, et al. (1991): "Selection Bias in TEFRA At–Risk HMOs". Medical Care 29: 318-331.

[19] Newhouse, J.P. (1996): "Reimbursing Health Plans and Health Providers: Efficiency in Production versus Selection." Journal of Economic Literature 34: 1236-1263.

[20] Pope, G.C., R.P. Ellis, A. Ash, et al. (2000): "Principal Inpatient Diagnostic Cost Group Model for Medicare Risk Adjustment." Health Care Financing Review, Spring 21(3): 93-118.

[21] Riley, G., J. Lubitz, and E. Rabey (1991): "Enrollee Health Status under Medicare Risk Contracts: An Analysis of Mortality Rates." HSR: Health Services Research 26(2): 137-163.

[22] Riley, G., C. Tudor, Y. Chiang, and M. Ingber (1996): "Health Status of Medicare Enrollees in HMOs and Fee-for-Service in 1994." Health Care Financing Review 17(4): 65-76.

[23] Shen, Y. and R.P. Ellis (2002a): "How Profitable is Risk Selection? A Comparison of Four Risk Adjustment Methods." Health Economics 11(2): 165-174.

[24] Shen, Y. and R.P. Ellis (2002b): "Cost-Minimizing Risk Adjustment." Journal of Health Economics 21: 515-530.

[25] Van Barneveld, E.M. et al. (2001): "Risk Sharing as a Supplement to Imperfect Capitation in Health Insurance: a Tradeoff Between Selection and Efficency." Journal of Health Economics 20: 147-168.

[26] Van Barneveld, E.M., L.M. Lamers, R.C.J.A. Van Vliet, and W.P.M.M. Van de Ven (1997): "Mandatory Pooling as a Supplement to Risk-Adjustment Capitation Payments in a Competitive Health Insurance Market." Social Science and Medicine 47: 223-232.

[27] Van Barneveld, E.M., R.C.J.A. Van Vliet, and W.P.M.M. Van de Ven (1996): "Mandatory High-Risk Pooling: An Approach to Reducing Incentives for Cream-Skimming." Inquiry 33: 133-143.

[28] Van de Ven, W.P.M.M., R.C.J.A. Van Vliet, E. Schut, and E.M. Van Barneveld (2000): "Access Coverage for High-Risks in a Competitive Individual Health Insurance Market: Via Premium Rate Restrictions or Risk-Adjustment Premium Subsidies?." Journal of Health Economics 19: 311-339.

[29] Van de Ven, W.P.M.M. and R.P. Ellis (2000): "Risk adjustment in Competitive Health Plan Markets." in: A.J. Culyer and J.P. Newhouse, eds., Handbook of Health Economics (Elsevier, Amsterdam) Chapter 14.

[30] Weiner, J.P., Dobson, A., Maxwell, S., et al. (1996): "Risk-Adjusted Medicare Capitation Rates Using Ambulatory and Inpatient Diagnoses." Health Care Financing Review 17: 77-100.

[31] Zhao Y, Ash AS, Ellis RP, Ayanian JZ, Pope GC, Bowen B, Weyuker L. (2005): "Predicting pharmacy costs and other medical costs using diagnoses and drug claims". Medical Care 43(1): 34–43.

Table 2: Descriptive statistics

Demographic Cell and Chronic illness	Mean	Std. Error
Cost	6944	21080
female 0-34	0.0056	0.0752
female 35-44	0.0103	0.1013
female 45-54	0.0134	0.1153
female 55-64	0.0207	0.1426
female 65-69	0.1138	0.3175
female 70-74	0.1359	0.3427
female 75-79	0.1149	0.3189
female 81-84	0.0848	0.2786
female 85-89	0.0514	0.2208
female 90-94	0.0231	0.1504
female over 94	0.0077	0.0877
male 0-34	0.0089	0.0942
male 35-44	0.0162	0.1265
male 45-54	0.0191	0.1371
male 55-64	0.0254	0.1573
male 65-69	0.0942	0.2921
male 70-74	0.1016	0.3022
male 75-79	0.0759	0.2648
male 81-84	0.0462	0.2100
male 85-89	0.0214	0.1447
male 90-94	0.0068	0.0825
male over 94	0.0016	0.0399
cancer	0.0112	0.1051
diabetes	0.1301	0.3364
CHF	0.0988	0.2983
stroke	0.0407	0.1978
COPD	0.1059	0.3077
renal	0.0129	0.1132

Table 3: Estimation of costs for acute and chronic care using only demographic information

Demographic Cell	Coefficient	Std.Error	t
female 0-34	4533	233.490	19.41
female 35-44	5262	173.003	30.42
female 45-54	6209	151.707	40.93
female 55-64	7396	122.172	60.54
female 65-69	4370	52.222	83.68
female 70-74	5433	47.783	113.70
female 75-79	6900	51.962	132.80
female 81-84	8490	60.471	140.40
female 85-89	10187	77.693	131.12
female 90-94	11314	115.743	97.75
female over 94	10624	199.899	53.15
male 0-34	4151	186.097	22.31
male 35-44	4810	138.112	34.83
male 45-54	5428	127.241	42.66
male 55-64	6666	110.561	60.29
male 65-69	5170	57.398	90.07
male 70-74	6455	55.250	116.83
male 75-79	8142	63.935	127.36
male 81-84	10321	81.912	126.00
male 85-89	12355	120.362	102.65
male 90-94	13658	212.755	64.20
male over 94	14114	441.036	32.00
number of observations	1417005		
R-squared	0.0103		

Table 4: Estimation of costs for acute and chronic care using demographic information for consumers with and without chronic illnesses

Demographic Cell	Coefficient	Std. Error	t
acute services for female 0-34	3737	242.00	15.44
acute services for female 35-44	4139	184.02	22.49
acute services for female 45-54	4153	171.70	24.19
acute services for female 55-64	4141	148.08	27.96
acute services for female 65-69	2682	58.52	45.83
acute services for female 70-74	3432	54.86	62.56
acute services for female 75-79	4364	61.20	71.30
acute services for female 81-84	5544	73.26	75.67
acute services for female 85-89	7257	96.95	74.85
acute services for female 90-94	8538	148.00	57.69
acute services for female over 94	8149	253.87	32.10
acute services for male 0-34	3458	190.23	18.18
acute services for male 35-44	3732	145.31	25.68
acute services for male 45-54	3372	142.43	23.68
acute services for male 55-64	3857	132.69	29.07
acute services for male 65-69	3320	65.51	50.68
acute services for male 70-74	4068	65.23	62.37
acute services for male 75-79	5012	78.01	64.24
acute services for male 81-84	6392	103.98	61.48
acute services for male 85-89	8135	156.43	52.01
acute services for male 90-94	9769	279.43	34.96
acute services for male over 94	9868	576.45	17.12
chronic services for female 0-34	8099	771.87	10.49
chronic services for female 35-44	7808	485.16	16.09
chronic services for female 45-54	8433	347.73	24.25
chronic services for female 55-64	9558	253.72	37.67
chronic services for female 65-69	7385	122.39	60.34
chronic services for female 70-74	7545	106.54	70.82
chronic services for female 75-79	8410	111.42	75.47
chronic services for female 81-84	8666	125.64	68.98
chronic services for female 85-89	7754	157.71	49.17
chronic services for female 90-94	6811	231.83	29.38
chronic services for female over 94	6196	401.69	15.42
chronic services for male 0-34	9491	704.09	13.48
chronic services for male 35-44	8628	411.08	20.99
chronic services for male 45-54	9059	299.00	30.30
chronic services for male 55-64	8576	231.87	36.99
chronic services for male 65-69	7210	129.36	55.73
chronic services for male 70-74	7826	118.11	66.26
chronic services for male 75-79	8960	131.98	67.89
chronic services for male 81-84	9851	164.64	59.84
chronic services for male 85-89	9893	239.52	41.30
chronic services for male 90-94	8869	422.00	21.02
chronic services for male over 94	9806	876.04	11.19
number of observations	1417005		
R-squared	0.0413		

Table 5: R-squared of the predictive models depending on the quality of the information

Model	$R-squared$
only demographic information or $\gamma = 0.50$	0.0103
demographic and chronic information $\gamma = 0.55$	0.0106
demographic and chronic information $\gamma = 0.60$	0.0114
demographic and chronic information $\gamma = 0.65$	0.0127
demographic and chronic information $\gamma = 0.70$	0.0146
demographic and chronic information $\gamma = 0.75$	0.0171
demographic and chronic information $\gamma = 0.80$	0.0203
demographic and chronic information $\gamma = 0.85$	0.0242
demographic and chronic information $\gamma = 0.90$	0.0290
demographic and chronic information $\gamma = 0.95$	0.0345
demographic and chronic information $\gamma = 1$	0.0414
number of observations	1417005

Table 6: Profits, enrollment and welfare loss under Minimum Welfare loss Risk Adjustment with high quality of information

δ_1	δ_2	α_{MWRAB}	β_{MWRAB}	Enrollees	Enr. ratio	profits	welfare loss
5	5	1.006	0.995	1297314	0.9115	452.70	5.161
5	10	1.008	0.995	1417005	1	0.06	5.267
5	15	1.039	0.995	1417005	1	4.51	7.243
5	20	1.071	0.995	1417005	1	2.59	9.283
10	5	1.006	0.995	1297314	0.9115	156.42	5.161
10	10	1.006	0.995	1297314	0.9115	310.48	5.161
10	15	1.007	0.995	1401195	0.9888	30.69	5.206
10	20	1.021	0.995	1417005	1	2.59	6.096
15	10	1.006	0.995	1297314	0.9115	14.19	5.161
15	15	1.006	0.995	1297314	0.9115	168.26	5.161
15	20	1.006	0.995	1297314	0.9115	322.32	5.161
20	20	1.006	0.995	1297314	0.9115	26.037	5.161

Table 7: Profits in millions, enrollment and welfare loss under all Risk Adjustment strategies

Payment System	α	β	Enrollees	Enr. ratio	Welfare Loss	Profits
NRA	1.010	0.992	1326682	0.9363	5.299	211.35
CRAA	1.009	0.993	1355768	0.9568	5.264	248.84
MWRAA	1.009	0.993	1355768	0.9568	5.264	217.55
CRAB	1.007	0.995	1401195	0.9888	5.206	74.49
MWRAB	1.006	0.995	1297314	0.9115	5.161	14.19

Table 8: Profits, enrollment and welfare loss under Outlier Risk Sharing

PP % of NRA	γ	Threshold	α_{ORS}	β_{ORS}	Enrollees	Enr. ratio	Welfare Loss	Profits
80	5,10	75,80,85	1.001	0.988	709955	0.501	4.95	21587
80	15,20	75	1.01	0.988	994271	0.7017	5.325	228*
80	15,20	80,85	1.001	0.988	709955	0.501	4.95	21587
85	5,10,15,20	75,80,85	1.01	0.988	994271	0.7017	5.325	3189*
90	5	75,80,85	1.01	0.988	994271	0.7017	5.325	11460*
90	5,10,15	75	1.011	0.988	1006832	0.7105	5.371	507*
90	5,10,15,20	85	1.01	0.988	994271	0.7017	5.325	10828
95	5,10,15,20	75,80,85	1.011	0.988	1006832	0.7105	5.371	4396*

* The value for the profits exposed are the minimum profits for all the combinations with the same enrollment, service distortion and welfare.

Table 9: Total welfare loss in millions for different levels of overprovision in the FFS sector

	WL0	WL5	WL10	WL15	WL20
NRA	323	530	981	1392	1746
CRAA	323	526	977	1389	1749
MWRAA	323	526	977	1389	1749
CRAB	326	520	971	1383	1756
MWRAB	302	516	967	1378	1722
ORS	120	495	946	1357	1445

Table 10: Welfare loss per HMO enrollee for different levels of overprovision in the FFS sector

	WL0	WL5	WL10	WL15	WL20
NRA	1113	1151	1555	1924	2262
CRAA	931	974	1327	1650	1945
MWRAA	931	974	1327	1650	1945
CRAB	566	593	806	1001	1179
MWRAB	452	470	635	785	923
ORS	504	505	668	816	952