Modest risk-sharing significantly reduces health plans' incentives for service distortion

SHULI BRAMMLI-GREENBERG*, JACOB GLAZER** AND RUTH WAITZBERG***.

Abstract

Public payers often use payment mechanisms as a way to improve the efficiency of the healthcare system. One source of inefficiency is service distortion (SD) in which health plans over/underprovide services in order to affect the mix of their enrollees. Using Israeli data, we apply a new measure of SD to show that a mixed payment scheme, with a modest level of cost sharing, yields a significant improvement over a pure risk-adjustment scheme. This observation implies that even though mixed systems induce overprovision of some services, their benefits far outweigh their costs.

Keywords

Service distortion, adverse selection, capitation, payment mechanisms, risk-adjustment, risk- sharing, managed care, managed competition.

* Myers-JDC Brookdale Institute and Haifa University, Israel. (e-mail: shuli@jdc.org).

**The University of Warwick and Faculty of Management, Tel Aviv University, Israel. (email: glazer@post.tau.ac.il);

*** Myers-JDC Brookdale Institute, Ben-Gurion University, Israel and Technische Universität in Berlin, Germany. (email: ruthw@jdc.org).

Acknowledgments. An earlier version of this paper was presented at the Risk-Adjustment Network Workshop held in October 2016 in Berlin, Germany. The authors are especially grateful to Konstantin Beck, Randall Ellis, Timothy Layton, Thomas McGuire, Amir Shmueli, Wynand Van de Ven, Richard Van Kleef, and other participants in the workshop for their helpful comments and advice. The authors are also grateful to Glied Sherry, Altman Stuart, Freed Gary and Wittenberg Raphael for their constructive suggestions. This study was funded by a grant provided by the Israel National Institute for Health Policy Research (grant number 2012/37). The authors declare no potential conflict of interest.

I. Introduction

Service distortion (SD) in healthcare occurs when health plans overprovide some services in order to attract more profitable enrollees and underprovide others in order to discourage less profitable ones.[1] SD is difficult to measure and regulate and therefore poses a major challenge to researchers and policy makers. In recent decades, they have been searching for efficient mechanisms to mitigate the incentive for SD among health plans and managed care organization (MCOs).[2] Since MCOs, which are paid primarily according to capitation, are rapidly becoming the dominant form of healthcare insurance and provision in many countries, the problem of SD is becoming a major issue of concern.

In addressing SD, policy makers face two major problems: (1) identifying those services that are more likely to be distorted; and (2) determining which mechanisms can be used in order to mitigate the incentive for SD without creating other types of inefficiencies. Building on the growing literature on service distortion, we will attempt to shed light on these two questions.

Using Israeli data, we compare the incentive for SD under several different reimbursement schemes using the *index of heterogeneity* (ψ), developed by Layton, Ellis and McGuire [1]. The ψ index (described in detail in Section III) is built on the assumption that a health plan allocates its resources across "budget cells" in a way that maximizes its profits, where the term "budget cell" refers to a set of services it provides, coupled with a group of enrollees who are the beneficiaries of those services. Examples might be: "Cancer treatment for women aged 25-34 living in Tel-Aviv" or "Pediatric services for children living in the

[1] Some papers use the term *service level selection* for what we refer to as *service distortion*. We use the terms "service distortion" and "incentive for service distortion" interchangeably in most of the paper except in places where the term "(incentives for) selection" seems more appropriate

[2] We use the terms Managed Care Organization and health plan interchangeably. An MCO is an entity that contracts to provide healthcare to its members who are often referred to as "subscribers" or "enrollees." In the US, there are several types of MCOs: Health Maintenance Organization (HMO), the Preferred Provider Organization (PPO), and the Point-of-Service plan. (Accountable Care Organizations (ACOs) can also be viewed as a type of MCO.) MCOs are also common in countries such as the Netherlands, Switzerland and Israel. MCOs have increasingly become the norm in the delivery of healthcare services over the past 15 years, partly due to the rising cost of healthcare. The commonly stated goal of MCOs is to reduce the cost of healthcare by negotiating lower fees for healthcare services, monitoring the types of treatment utilized by subscribers, and reviewing the relative cost effectiveness of treatments.

North." By allocating resources across budget cells, the health plan determines the quantity and quality of the various services it provides to each group, which, in turn, determine the distribution of individuals who choose to join or leave the health plan. If, for example, the plan predicts that the expected profit on a woman aged 25-34 living in Tel-Aviv and suffering from cancer will be negative whereas the expected profit on a healthy child, living in the North, will be positive, then the plan has the incentive to allocate more resources to the latter's budget cell and less to the former's. The ψ index enables us to compare different reimbursement mechanisms vis-à-vis their effect on this incentive.

The aim is to show that a mixed payment scheme, with a modest level of risk-sharing, yields a significant improvement over a pure risk-adjustment scheme.

We analyze empirically two forms of risk-sharing between health plans and the payer in a managed competition system with risk-adjusted capitation payments. It is assumed that the purpose of risk-sharing is to reduce the health plan's incentives for selection while maintaining their incentives for efficiency to whatever extent possible (i.e. minimizing "insurer moral hazard").

In the first part of the analysis, it is shown that under an improved risk-adjustment reimbursement scheme, health plans still have a relatively strong incentive for SD. We then consider the incentive for SD under two mixed reimbursement schemes, in which the payment to a health plan is primarily determined according to prospective capitation, although a small degree of risk-sharing is also incorporated. In order not to excessively increase the incentive for moral hazard, we incorporate only a very modest level of risk-sharing in each of the two schemes. The first scheme compensates the health plan retrospectively for 25% of the actual expenditures on the costliest percentile of its members, while the remaining 75% (and 100% of the expenditure on the remaining 99% of the enrollees) is reimbursed prospectively on the basis of the refined risk-adjustment scheme mentioned above. We show that even such a small modification of the reimbursement scheme dramatically reduces a plan's incentive for SD. Under the second risk-sharing mechanism, the health plans are compensated retrospectively for 12.5% of their total hospitalization costs (which comprise about 5% of their total costs, on average) and the remaining compensation is paid prospectively according to the refined risk-

adjustment scheme described above. The latter mechanism is the most efficient of those considered in reducing incentives for selection.

It is important to stress that the objective is not to evaluate the performance of the current Israeli risk-adjustment formula (which is certainly far from optimal) but rather to use the Israeli data to demonstrate how a modest level of risk-sharing can significantly improve upon any pure risk-adjustment payment scheme. We believe that our main insights are relevant for any healthcare system in which managed care organizations compete for enrollees and are reimbursed by a third party (usually prospectively and by capitation). The results show that even a small amount of risk-sharing can significantly lower the incentive for SD. Clearly, risk-sharing mechanisms create a variety of new distortions; nevertheless, and especially given the very small weight we attribute to this mechanism in the payment formula, they are likely to be of a second order of magnitude relative to the benefits they deliver.

II. Background

In the first part of this section, we examine the literature on SD which is motivated by the incentive for selection. The leading assumption in this literature is that the premium is heavily regulated and often not even paid directly by the enrollees. In the second part, we discuss the literature that compares between the various payment mechanisms meant to address the selection problem.

2.1 Service distortion

In general, adverse selection arises because consumers possess (private) information that is not accessible to the insurers/health plans or because the insurer is prohibited from conditioning insurance premiums on observable information like age, gender, and medical status.

Two types of conceptual frameworks are usually used to capture selection. The first is the *fixed contracts* approach, which takes insurance contract provisions as given and views selection as influencing only insurance prices in equilibrium. The second is the *endogenous contracts* approach, in which selection also affects the design of the contract itself, including the overall level of coverage and coverage for services that are differentially demanded according to the health of the consumer [2]. In a managed care setting, premiums

that health plans can charge their enrollees are usually regulated. However, a plan can use observable information even in a fixed premiums contract setting by determining service quality or quantity in order to affect the type of enrollees that will join it. SD is therefore a result of the incentive to influence enrollee types by over- or undersupplying certain health care services [3].

In spite of the loss in welfare involved, little attention has been paid to SD in the literature until recently. Part of the explanation may be the difficulty in identifying SD directly since it requires identifying the efficient level of care for each individual and then measuring the actual level of care received. It is not surprising therefore that most of the literature on SD has focused on the *incentives* for distortion in the different systems, rather than measuring the distortion itself.

Several papers have suggested methods to identify services that are more likely to be distorted or to be rationed tightly by health plans.[3] The pioneering paper by Glazer and McGuire [4] proposes and derives formulas for optimal risk-adjustment payments to health plans that mitigate incentives for SD. Frank, Glazer, and McGuire [5] assume that health plans ration care by setting a "shadow" price for each service, such that each enrollee receives treatment up to the level at which his marginal utility from a particular service equals the shadow price of that service. Using the shadow price approach, they identify some characteristics of the services that are more likely to be over- or underprovided. One of their main insights is that plans have the incentive to spend more on services used primarily by more profitable enrollees and less on services used primarily by unprofitable ones (such that there is a positive correlation between the cost of a service and the profitability of the enrollees that use it). Building on their work, Ellis and McGuire [6] show that the services most likely to be distorted by the health plans with the goal of attracting profitable enrollees are characterized by: (1) a high variance in utilization of the service among enrollees;)2) high *predictability* (how well spending on certain services can be anticipated by the individual); and)3) high *predictiveness* (i.e. how closely the predicted levels of certain services contemporaneously co-vary with total healthcare spending). Based on the selection index presented in Ellis and McGuire [6] and using disaggregated

[3] Under the assumption of profit maximization.

commercial insurance claims, Ellis, Jiang and Kuo [7] show that managed care plans spend proportionally less on those types of services that are predicted to be more profitable if rationed tightly.

McGuire et al. [8] improve the selection index based on predictability / predictiveness by taking into account the role of premiums in financing plans. Moreover, by incorporating demand-side cost-sharing in their analysis and relying on an explicit model of a plan's profit maximization, they are able to measure incentives based on the predictability and predictiveness of various medical diagnoses. They found that among the chronic diseases studied, health plans have the greatest incentive to skimp on care for cancer, mental health and substance abuse. Ellis, Martins and Zhu [3] refine McGuire et al. [8]'s index and calculate the full selection elasticity, which reflects five elements apart from predictability and predictiveness: 1. demand-side cost-sharing; 2. service-level demand elasticities; 3. individual-level profit variation; 4. actual profit levels; and 5. demand responsiveness of health insurance enrollment to expected spending. Their analysis confirms the findings of Ellis and McGuire [6] and of McGuire et al. [8] that selection incentives are stronger for services that tend to be provided by non-managed care plans than for those that tend to be provided by managed care plans. They also found that incorporating service-level demand elasticities into the analysis significantly changes the conclusion regarding the selection incentives for some services.

Geruso and Layton [2] discuss the policy tools commonly used to address the problem of selection in a competitive insurance market. Geruso, Layton, and Prinz [9] is one of the few empirical papers on non-price contract distortions. They use the Exchange regulator's risk-adjustment and reinsurance algorithms to simulate enrollee-specific net revenue based on data for prescription drug expenditure and diagnoses. They compare the simulated Exchange revenues to the directly observed claims costs, yielding estimates of person-specific implied profits. To understand the potential for screening unprofitable patient types on the basis of drug coverage, they group patients by prescription drug consumption into various therapeutic classes. Our paper adds also to this small body of literature by comparing a health plan's revenues to its costs. We also specifically model the mechanism by which health plans allocate resources across sets of services and groups of enrollees. Following Layton, Ellis and McGuire [1], and unlike most of the earlier research that has

(sometimes implicitly) assumed that plans allocate resources across specific services or individuals, we assume that plans allocate resources across "budget cells" that consist of both (sets of) services and (groups of) individuals for whom these services are intended. The idea that health plans make their decisions on the basis of budget cell allocation (i.e. at the group-services level) rather than at the individual/type or the specific service level is, in our opinion, more in line with the actual decision-making process of health plan managers. Using the budget cells concept, the same service can appear in multiple budget cells if it is provided to various groups of individuals, while the same individual can appear in multiple budget cells if he uses a number of different services. This structure enables us to pinpoint which enrollee groups are more subject to SD, as well as which services are more likely to be distorted.

2.2 Payment mechanisms as a tool to mitigate a health plan's incentive for service distortion

The main policy tool for reducing a health plan's incentive for SD is the design of an appropriate payment mechanism. The question of how a payment mechanism can be used to mitigate this type of distortion has been widely discussed in the literature [4, 10-15]. One of the key insights is that the payment mechanism is usually more effective in achieving its goals if it is a mixture of prospective and retrospective reimbursement schemes, thus trading off incentives for selection and *moral hazard* [4, 12-14,16-19].

In most managed care systems, health plans are usually paid according to a risk-adjusted capitation formula [20], which involves a prospective payment mechanism that pays health plans per enrollee (prior to use) for a predetermined period of time (usually a year). The regulator sets a per capita payment based on a relative risk scale determined for sub-groups in the population. In other words, the health plan receives a fixed sum times the number of enrollees in each group defined by the capitation formula. Many countries attempt to reduce the incentive for selection by improving the risk-adjustment component of the capitation formula [21]. The more fine-tuned the risk-adjustment component, the less incentive there will be for distortion [22].

Nonetheless, many studies have shown that the ability of a capitation formula to explain and predict the expected expenditure and variance in the consumption of health services is limited. Thus, demographic and socioeconomic variables predict only about 10% of the

variance, while adding additional background variables and health status raises the explained variance to only about 20% [6,12]. In other words, most advanced risk-adjustment mechanisms using prospective (i.e. predetermined) information do not fully explain the variation in expected expenditure and do not completely eliminate a plan's incentive for selection [23].

The standard tool used to measure a payment mechanism's incentive for adverse selection, and in particular SD, is statistical: The R-squared from a regression of actual costs on risk adjustor variables. Other measures include the "predictive ratios", i.e. the ratio of predicted costs to actual costs for selected groups in the population, such as the chronically ill, and "over- and undercompensation", i.e. the difference (rather than the ratio) between projected revenues and costs [1, 24-25].

The use of R-squared as a measure of incentives for selection implicitly assumes that plans can discriminate among enrollees on an individual basis. This is a strong assumption and to a large extent unrealistic, especially when selection is achieved via services rather than premiums. On the other hand, it is unlikely that SD takes place solely at the service level and that it is the same for all individuals, an assumption implicitly made by some studies of service selection. The approach adopted here represents what we believe to be a more realistic mixture of these two extremes. Thus, we assume that plans allocate resources across budget cells and thus discriminate among groups of individuals and groups of services. As a result, two different individuals with the same medical problem and using the same service may receive different levels of that service if they do not belong to the same budget cell. Similarly, the same individual may receive a high level of treatment for one medical problem but a low level for another, if the two do not belong to the same budget cell.

In the analysis, we assume that the budget cells are given and examine how plans allocate resources among them. A more general analysis would also look at how plans decide on their budget cells and how such a decision is affected by (among other things) the payment mechanisms by which the plan is reimbursed.

III. Measuring incentives for service distortion (SD)

We employ the ψ index, developed by Layton, Ellis and McGuire [1],[4] to study the incentive for selection under various reimbursement mechanisms, rather than the traditional R-squared measure.

Like the "payment system fit" measure presented by Layton, Ellis, McGuire and Van Kleef [25], our measure evaluates payment mechanism alternatives according to the incentive for distortion they create. Moreover, it is also able to identify which enrollee groups are more susceptible to SD, as well as which services are more likely to be distorted.

In order to employ the measure, we need to identify the health plan's "budget cells". To do so, it is assumed that plans allocate resources across budget cells in a way that will make them attractive to some individuals (the more profitable ones) and unattractive to others (the less profitable ones). If compensation is not optimal in the sense that it does not fully compensate the health plan for the expected cost of each enrollee, it will have an incentive to "over-allocate" resources to budget cells that are more profitable and to "under-allocate" to budget cells that are less so.

The measure of SD is built on a result obtained in Frank, Glazer and McGuire [5]. They study a model of imperfect competition among health plans, where each of them chooses (only) the quality of the different services it provides. Quality of service s is modeled as a "shadow price", λ_s, such that each enrollee receives treatment up to the level where her marginal utility from that treatment equals the shadow price. In other words, if λ_s is the shadow price of service s and $u_s^i(x_s^i)$ is individual i's utility from receiving x_s^i units of service s (where x_s^i is measured in dollars), then x_s^i is determined by $u_s^{i\prime}(x_s^i) = \lambda_s$. Thus, the higher the shadow price that a plan chooses for service s, the lower the level of that service provided to all enrollees. Frank, Glazer and McGuire [5] show that in equilibrium the (profit-maximizing) plan will choose the shadow price in such a way that $\frac{1}{\lambda_s} \approx \sum_i \frac{x_s^i \pi^i}{x_s^i}$.

[4] We are using the version of the index developed in Layton, Ellis and McGuire [1] rather than the final formula appearing in Layton, Ellis, McGuire and Van Kleef [25]. The former is based on group-level deviations while the latter is based on individual-level deviations. We chose the group-level formula since we believe it to be a more realistic description of how health plan managers implement their desired policy. We are not aware of any previous research that has looked at how health plan managers actually implement service distortion – a question left for future research.

This implies that the quality of service s decreases (λ_s increases) with x_s (the average cost of providing the service) and increases with $\sum_i x_s^i \pi^i$ (the correlation between enrollees' use of service s and the profitability of that use to the health plan). Intuitively, the higher the average cost of providing a service, the weaker will be a health plan's incentive to provide it. On the other hand, if individuals who use more of a particular service are generally more profitable to a health plan, then there will be a strong incentive to provide a high level of that service.

We utilize the result obtained in Frank, Glazer and McGuire [5], though we focus on budget cells rather than services. We assume that a plan's incentive to allocate resources to a given budget cell decreases with the average cost of all services in that cell and increases with the profitability of the individuals who use more of the services in that cell.

The index of heterogeneity

The ψ index is defined as follows:[5]

(1)

$$\psi = 1 - \frac{\sum_q \left(\sum_i \frac{x_i^q}{x^q} (rev_{p,i} - x_i)\right)^2}{\sum_q \left(\sum_i \frac{x_i^q}{x^q} (\bar{x} - x_i)\right)^2}$$

where:

- q is a budget cell within a matrix Q of 96 cells that we define according to:
  - Four medical service categories denoted by s (ambulatory services in the community, hospitalization, use of medication and other medical services).
  - Three age groups: 0-24, 25-64 and 65 and older.
  - Gender
  - Four areas of residence (North, Tel Aviv area, Jerusalem area, and South).

[5] This is the index of heterogeneity (ψ) developed by Layton, Ellis and McGuire [1] and mentioned above.

Examples of a budget cell would be: ambulatory services for men aged 25-64 who live in Jerusalem or hospital services for women aged 65 or older who live in Tel Aviv.

- x_i is the plan's total expenditure on individual i while $rev_{p,i}$ is the revenue received by the health plan for individual i under payment mechanism p. A health plan's profit on individual i is equal to revenues minus costs, namely $(rev_{p,i} - x_i)$ and $\bar{x}$ is the average expenditure per enrollee.
- x_i^q is the health plan's total expenditure on individual i in the budget cell q, and x^q is the plan's total expenditure in cell q (on all services and all individuals in the cell). Thus, the ratio $\frac{x_i^q}{x^q}$ is individual i's share of the plan's spending on budget cell q. We assume that if a plan allocates one additional dollar to budget cell q and individual i belongs to that budget cell, her use of all services in that budget cell combined will increase by $\frac{x_i^q}{x^q}$, namely in proportion to the individual's current use of those services. Hence, the ratio $\frac{x_i^q}{x^q}$ can be thought of as the measure of individual i's incentive to join the plan as a result of an increase in the budget allocated to the cell q. The term $\left(\sum_i \frac{x_i^q}{x^q}(rev_{p,i} - x_i)\right)^2$ can therefore be thought of as a measure of the plan's incentive to allocate more funds to the budget cell q, under the payment mechanism p.

The health plan's incentives are summed over all q cells to fully characterize the distortion incentive under a given payment mechanism.

The value of ψ ranges from 0 (strongest incentive for SD) to 1 (no incentive for SD) and a change in the payment mechanism that results in an increase in ψ decreases the incentive for SD.

IV. The Data and the Empirical Analysis

The main objective is to calculate the ψ index under various possible payment mechanisms, using Israeli data. In Israel, the health plans are almost fully funded from the public purse, according to a risk-adjusted capitation formula. Therefore, when calculating the ψ index

for the current Israeli capitation mechanism, the revenue weights ($rev_{IsreliP,i}$) are taken to be the risk-adjusted Israeli capitation rates (Table 1).[6]

Table 1: The current Capitation Formula Rates[7] (1=overall mean)

	2010 Capitation rates			
	Center		Periphery	
AGE	Female	Male	Female	Male
Newborn	1.41	1.87	1.45	1.92
1-4	0.75	0.94	0.80	0.99
5-14	0.38	0.41	0.42	0.45
15-24	0.43	0.36	0.47	0.40
25-34	0.73	0.41	0.77	0.46
35-44	0.78	0.57	0.82	0.62
45-54	1.14	0.99	1.18	1.03
55-64	1.70	1.79	1.74	1.84
65-74	2.63	3.14	2.67	3.18
75-84	3.40	4.13	3.45	4.18
85+	3.52	4.23	3.57	4.27

For example, for a woman aged 40 living in the Center, the plan receives compensation equal to 0.78 of the average compensation (i.e., $R_{j=(35\text{-}44,\ female,\ Center)} = 0.78$).

When calculating the ψ index for the three other mechanisms examined in the paper, the revenue weights ($rev_{P,i}$) are calculated accordingly. It is important to stress that in all four mechanisms, $\bar{x}$ was either 1 or normalized to 1 in order to obtain the same grand mean. This was done in order to avoid downward bias in the calculation of the index.

4.1 Data

The data are taken from the 2009 National Health Survey carried out by the Israeli Central Bureau of Statistics (CBS).[8] This is the same basic data set used by the Israeli Capitation Committee to calculate the 2010 capitation formula rates. The survey population included permanent residents of Israel, as well as tourists and temporary residents living in Israel continuously for more than one year. The main goal of the survey was to provide data on

[6] For a detailed description of the Israeli payment mechanism, see Appendix 1.

[7] The current Israeli capitation formula was last updated in 2010. Table 1 presents the most recent rates.

[8] For the distribution of the 2009 health survey population by age, gender and place of residence, as well as the distribution of respondents with a chronic illness, disability, depression or anxiety or cancer by age and gender, see Appendix 2.

health status, health service utilization, health habits and health insurance. The survey did not include individuals living in institutions or in non-recognized settlements. The response rate was 82% and a total of 8,728 households (consisting of 28,968 individuals) were surveyed. The survey was conducted by phone or face-to-face when phone interviews were not possible. The survey questionnaire was also available in Arabic and Russian. The survey data were weighted according to CBS conventions.

Data on the actual expenditure of each individual is not available at the national level, and therefore we have used the data on individuals' use of services (measured in "physical" quantities of each service) to calculate expenditure. Focusing on the use of health services is potentially problematic since it may not result in an accurate estimate of a health plan's costs. However, the weights we assign to the different services in order to estimate costs are the same as those used by the Israeli regulator and since our main objective is primarily to compare the incentives for selection under different payment mechanisms, we believe that the relative cost estimates are sufficiently accurate for our purposes. Furthermore, the relative costs we obtain (as a function of enrollee type) look very similar to those often reported in the literature.

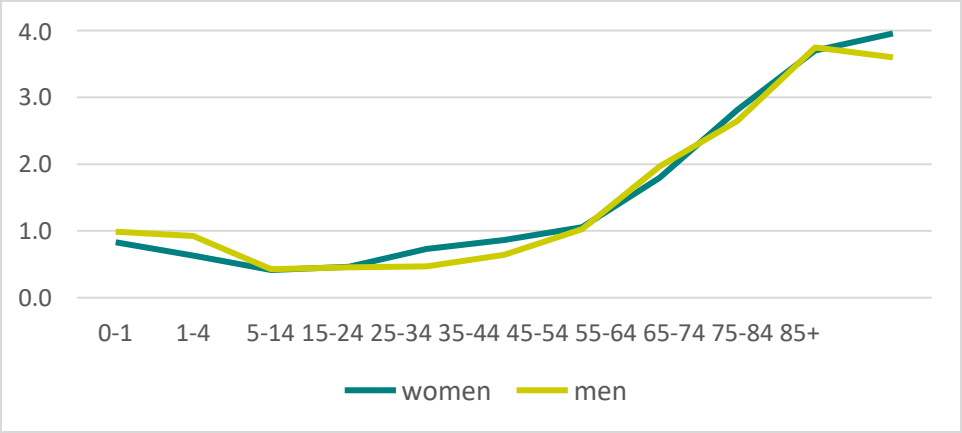
Another limitation arises from the use of survey responses to estimate utilization. First, the survey data are self-reported and therefore subject to recall bias and an individual's willingness to report the truth. Second, the survey used categories of healthcare services that overly aggregate the responses for the purpose of estimating expenditure. For example, the cost of "hospitalization" services varies significantly from one type of hospitalization to another; the cost of a visit to a GP differs from that of a visit to a specialist; and the cost of medication varies to an even greater degree, such that the data provide only approximations of the actual costs incurred by the health plans. Finally, individuals living in institutions (nursing homes, hospices or extended hospitalization) were not included in the survey data, though they contribute significantly to a health plan's expenditure.

4.2 Results

Actual cost-of-use rates

Based on the distribution of use for each individual per service category (*s*) and Equation (A.2) (see Appendix 1), the actual use rates were calculated according to *j* groups.[9] Figure 1 presents the total actual use rates by age and gender. Not surprisingly, there are differences between men and women in the actual use of services, such that men use more healthcare services than women in the early age group (1-14) while women use more during the period of fertility (15-54) and in the elderly age group (85+).

Figure 1: Actual rates of health service use, by age and gender



In Appendix 3, we present the actual use rates by service category for men and women by age. The results show that the rates for outpatient services in the 15-54 age groups are significantly higher for women than for men. For both men and women, there is a steep rise in the rate of inpatient care from age 45 onward.

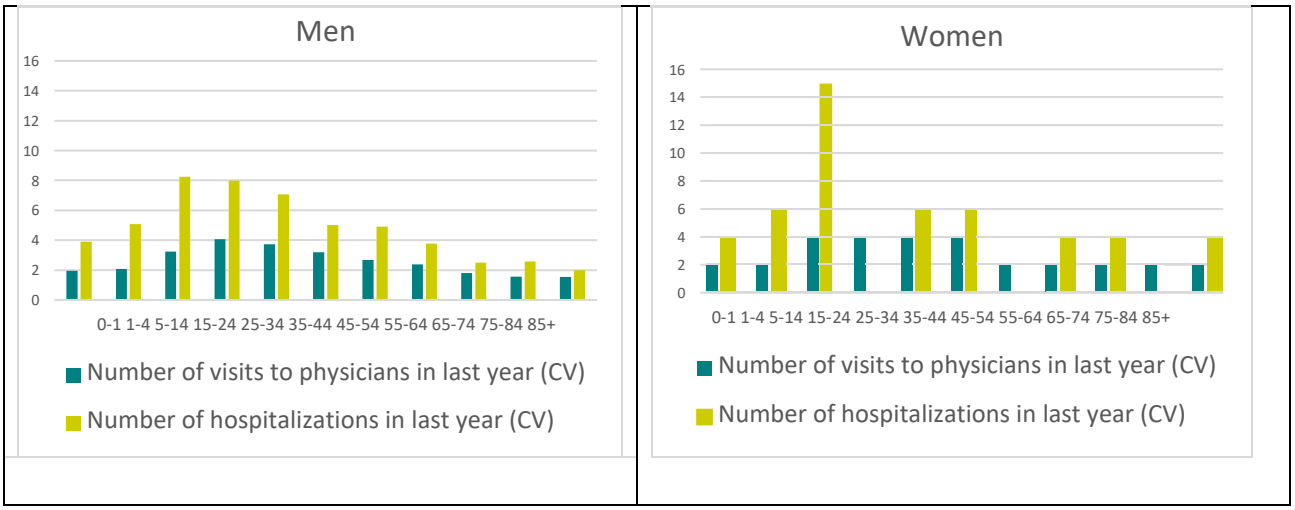
Using the index presented in Appendix 3, we examine which groups are preferred more by a health plan and which are preferred less. No capitation formula, regardless of how fine-tuned it is, can match premiums to the exact risk of each enrollee. Therefore, within any capitation group, there will always be some enrollees who are preferred over others (e.g. healthy ones over sick ones). In order to explore which groups of enrollees are driving the incentives to distort services, we chose to use the CV (coefficient of variation) index. This

[9] In order to simplify the analysis, we do not include residence in the periphery as an adjuster in the models. However, this should not affect the results since the periphery adjuster only adds a fixed (and relatively small) sum to the payment for all enrollees living in the periphery regardless of age or gender.

index enables us to assess the incentives for selection within each group under the current mechanism. It represents the dispersion of costs within a group and is calculated by dividing the standard deviation of the actual use of each service by the average. Because of the adjustment to the average, the index allows for a comparison between groups and between variables.

Figure 2 shows that the variance within groups, where the largest variance is found among girls aged 5-14, although the index is significant in other groups as well.

Figure 2: Coefficient of variation (CV) within each group in the actual use model by age and gender



Note that the less advanced the capitation model, the larger will be the variance within each group.

Predicted cost-of-use rates

In this section, we derive an "improved" risk-adjustment reimbursement scheme. It is shown that even under this scheme, health plans have an incentive for SD and therefore other payment tools should also be considered.

The improved risk-adjustment scheme includes variables that might not be considered by the payers to be feasible or available risk adjusters, though they may be relevant when the health plans make their decisions regarding the provision of services. In order to carry out the most profitable “selection strategy”, health plans need to predict, as accurately as possible, each enrollee’s expected revenue and expected cost. As for expected cost, we assume that the health plans use all the data available to construct a model of predicted use.

Therefore, we estimated the predicted use of services given explanatory variables that are available to the health plans (but not necessarily to the payer). Variables that are known only to the enrollee but not to the health plans (such as subjective physical and mental health status) were not included in the model even though they have strong explanatory power in an equation for use of health services. The predicted use was calculated from a two-stage regression model, where the first stage estimated a logistic regression of the probability of using a service category (as described above) while the second estimated a generalized linear regression for the amount of use (if there was any). The predicted cost-of-use rates were also calculated based on Equation (A.2). In what follows, the model that best estimates the predicted cost of use will be referred to as the "**improved" capitation formula** (i.e., the case in which the health plans are paid according to the best predicted cost of use).

Each regression model was chosen based on a robustness analysis and included covariates that significantly improved the goodness of fit. These included health status, disability and demographic, socioeconomic and health behavior variables. The results of the multivariate analysis are shown in Tables A.5 to A.9 in Appendix 2.

The explanatory variables with the strongest independent effect (highest Odds Ratio and highest Wald chi-square) on both the *probability* of using healthcare services and the *amount* of services used are chronic illness and physical disability. Having cancer or suffering from depression significantly affects only the probability of using health services.

Moreover, the two-stage regression model also shows that engaging in physical activity had a negative correlation with the probability of being hospitalized, the number of hospitalizations per year, and the use of medication; on the other hand, it had a positive effect on the probability of visiting a physician and related health professionals. Smoking was found to be positively correlated with the probability of being hospitalized. Arabs had a higher probability of being hospitalized and using medication. Finally, individuals living in areas characterized by a low socioeconomic level had a lower probability of using services.

V. The ψ index for four different payment mechanisms

The ψ index enables us to compare various payment mechanisms in terms of their incentives for SD. We will apply the index in order to compare the incentives under the current payment mechanism in Israel to those under alternative mechanisms. Note that the risk-adjustment formula currently applied under the Israel NHI Law is quite "simple" since it uses essentially only three risk adjusters. Initially, we compare the incentive for distortion under the current system to the incentive under the improved risk-adjustment formula, which includes additional risk adjusters. We then explore how the introduction of risk-sharing mechanisms affects the incentive for SD. We chose to analyze two risk-sharing mechanisms: one that focuses on high risk individuals and another that focuses on high cost individuals. Both groups are characterized by relatively high inelasticity of demand and hence by low welfare loss generated by moral hazard. Thus, our mechanism seems to be able to mitigate the service distortion problem without overly exacerbating the moral hazard problem.

5.1 Comparison of the ψ index under the current capitation formula to that under the improved capitation formula

Figure 3 depicts the value of the ψ index for the two capitation mechanisms (*p*): the current Israeli risk-adjustment scheme (the "current" capitation mechanism) and the improved risk-adjustment formula (the "improved" capitation mechanism).

The ψ index is 0.11 higher under the improved capitation scheme than under the current capitation scheme (0.25 versus 0.14)]. Nonetheless, an index of 0.25 suggests that the intragroup incentive for SD remains even when a more refined capitation model is implemented.

5.2 The ψ index for payment mechanisms that include risk-sharing

Risk-sharing mechanisms (RSMs) reduce the incentive for selection. They do not usually replace the capitation mechanism but rather complement it. We calculated the ψ index for two compensation mechanisms that combine both prospective payments (capitation) and RSM.

The first RSM compensates the health plans retrospectively for 25% of the actual expenditure on the highest percentile of individuals by cost. The remaining 75% of their

costs are reimbursed prospectively (according to the "improved" capitation formula described above). Thus, the health plan's revenues for the highest cost percentile are:

$$rev_{1\%,i} = 0.75 * rev_{predicted,i} + 0.25 * xi;$$

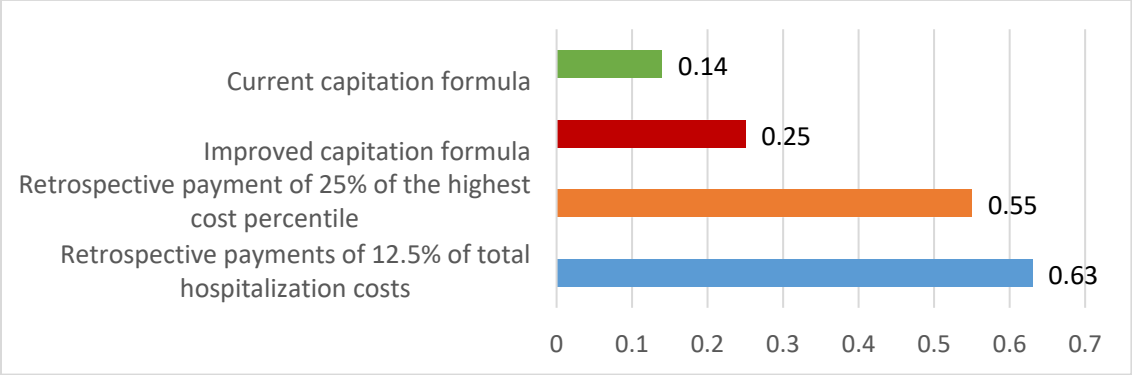
and for all other individuals are equal to $rev_{predicted,i}$.[10,11]

The second type of RSM incorporates a separate payment by type of treatment. Since in our case hospitalization is the largest category of expenditure for the health plans ($W_{hospitalization}$=0.4), we chose it as the treatment to be partially covered by the payer. More specifically, health plans are compensated retrospectively for 12.5% of total hospitalization costs.[12] In this case, the health plan's revenues are:

$$rev_{P,i} = 0.95\, rev_{predicted,i} + 0.125 x_i^{hospital}$$

The ψ indexes for the two risk-sharing payment mechanisms are shown in Figure 3 alongside those for the current capitation and the improved capitation mechanisms.

Figure 3: **ψ** index for the four compensation mechanisms



As shown in Figure 3, the ψ index is much higher for the combined prospective-retrospective mechanisms than the best prospective payment mechanism, suggesting a lower incentive for SD under the combined payment mechanisms. The payment

[10] Not surprisingly, the top percentile of individuals according to cost accounts for 22% of the total budget; therefore, 25% of the actual expenditure on them is equal to about 5% of the total budget.

[11] In order to keep the total budget identical across all the examined payment mechanisms, the average revenue per individual is normalized to 1 for all of them.

[12] A retrospective payment covering 12.5% of hospitalization costs represents 5% of a health plan's actual expenditure (since the weight of inpatient services is 0.4).

mechanism in which the payer assumes some of the health plan's risk for expensive treatments (in this case hospitalization) appears to be more effective in reducing the incentive for SD. The disadvantage of this type of mechanism is that it may encourage moral hazard or the provision of unsuitable treatment in order to receive payment (supply-induced demand). However, we incorporate only a very modest level of risk-sharing in each of the two schemes, i.e. only 5% of the health plan's expected total costs. This is a very small change in the payment scheme and therefore will induce only a very small change in power, as defined by Geruso and McGuire [26]. Given the very low level of risk-sharing, supply-induced demand and insurer moral hazard should not be a major concern.

VI. Conclusions and Discussion

We compare four different payment mechanisms using the ψ index and Israeli data in order to determine which achieves the lowest incentive for SD. The estimation results can also be used to identify the enrollee groups most susceptible to SD and which services are more likely to be distorted.

The index uses the concepts of "budget cells" and profit per enrollee, which add another dimension to the techniques for estimating SD incentives that already appear in the literature. In our setup, health plans allocate resources across budget cells so as to attract more profitable enrollees and discourage less profitable ones, based on the "predicted cost" and "predicted revenue" for each enrollee.

We assume that "predicted costs" are more relevant than actual costs in the considerations of the health plans. This could also be referred to as "ex ante adverse selection". Enrollees included in unprofitable "cells" are more exposed to the distortion of a service they are more likely to use. However, health plans have an incentive to distort not just one specific service used by the enrollees in this unprofitable cell; but other services they may use as well. For example, we found in the Israeli data that individuals who are more likely to be hospitalized are more exposed not only to hospitalization distortion but also to distortion of other services. Therefore, improving the payment mechanism by including retrospective payment for hospitalization costs, in addition to the risk-adjusted capitation component, led to the greatest overall reduction in the incentive for SD. The demand for hospitalization

is relatively inelastic and therefore a modest retrospective payment is not likely to have a major effect on moral hazard.

The health plans have a variety of tools to carry out SD. For example, they can contract with fewer specialists / hospitals in a specific location and in that way limit the choice possibilities of their enrollees in that location. By contracting with physicians with specific characteristics (such as gender or native language) or by "playing around" with reception hours throughout the week (i.e. mornings vs. evenings, weekends) they can deter or attract specific groups of enrollees. An example for this is the numbers published on May 2019 – although almost one and a quarter million women in Israel are at age of menopause, there are only about three to four dozen physicians specializing in the menopause's symptoms, and who are qualified to treat it. Other examples include the requirement of obtaining administrative approvals for certain medical procedures, tests and medications; limiting the resources transferred by the health plan to a particular hospital; and using waiting time as a shadow price.

The main payment mechanism used to reimburse health plans in Israel is a capitation formula with a relatively simple risk-adjustment mechanism (relative to those used in other countries). The formula has been updated only four times during the past twenty years, which exacerbates the gap between payments and the actual cost to the health plans and strengthens the incentive for selection.

Due to the complex nature of healthcare provision, it is difficult to present evidence of risk selection. In the Israeli system, there is some limited evidence, summarized by Brammli-Greenberg, Glazer and Shmueli [27], that the aforementioned incentives do indeed exist and that they lead to *ex-ante* individual selection and SD. For example, a study that was based on the actual costs of the largest Israeli health plan (Clalit) showed that in 2003 compensation by the capitation formula for chronically ill enrollees was, on average, 60% lower than the plan's actual expenditure. For diabetic patients, the level of under-compensation reached 71%. Conversely, Clalit was about 60% over-compensated, on average, for members without any chronic illness. A more recent study by Shmueli and Nissan-Engelcin [28] found that a health plan's costs are higher for individuals living in wealthy or more central locations. Since the rich enjoy better health, this effect might

reflect the lower supply of (ambulatory) health services in poor and remote locations. Another study found that the availability of physicians' services (measured as reception hours per age-adjusted inhabitant) in areas with low levels of health (as measured by the Standardized Mortality Rate) and in poor areas is significantly lower than in areas with higher levels of health.

Some countries use risk adjusters to achieve healthcare policy goals. For example, in the Netherlands risk adjusters are used to promote the treatment of diabetes and to encourage the health plans to attract diabetics [21]. However, including diagnosis as a risk adjuster may have negative impacts, such as incentives for up-coding, over-diagnosis or overtreatment. Moreover, it would appear that using a capitation formula as a tool to promote policy is a complex task, since constructing the capitation formula depends on (and therefore is also limited by) the availability of data to calculate the risk adjusters and their scales. International experience indicates that combining complementary payment mechanisms with capitation is the best-practice mechanism for carrying out health policy and reducing incentives for selection.

Although the analysis is based on survey data, a strong and robust reduction in the incentive for SD was found when the current capitation model is replaced by the best predicted-capitation formula and an even greater reduction was achieved when it is replaced by a payment mechanism that combines the best predicted-capitation with a modest RSM component, such as reinsurance or retrospective payment.

I. References

1. Layton, T. J., Ellis, R. P., and McGuire, T. G. 2015. "Assessing incentives for adverse selection in health plan payment systems", *National Bureau of Economic Research* (No. w21531).
2. Geruso, M., & Layton, T. J. 2017. "Selection in Health Insurance Markets and Its Policy Remedies". *Journal of Economic Perspectives*, *31*(4), 23-50.
3. Ellis, R. P., Martins, B., and Zhu, W. 2017. "Demand elasticities and service selection incentives among competing private health plans." *Journal of Health Economics*, 56, 352-367.
4. Glazer, J. and McGuire, T.G. 2000. "Optimal risk-adjustment of health insurance premiums: an application to managed care". *American Economic Review*, 90(4): 1055-1071.
5. Frank, R.G.; Glazer, J. and McGuire, T.G. 2000. "Measuring adverse selection in managed health care". *Journal of Health Economics*, 19(6): 829-854.
6. Ellis, R.P. and McGuire, T.G. 2007. "Predictability and Predictiveness in Health Care Spending". *Journal of Health Economics* 26 (1): 25-48
7. Ellis, Randall P., Shenyi Jiang, and Tzu-Chun Kuo. 2013. "Does service-level spending show evidence of selection across health plan types?." *Applied Economics* 45, no. 13: 1701-1712.
8. McGuire, T.G., Newhouse, J.P., Normand, S.L., Shi, J. and Zuvekas, S., 2014. Assessing incentives for service-level selection in private health insurance exchanges. *Journal of Health Economics*, *35*, pp.47-63.
9. Geruso, M., Layton, T., & Prinz, D. (2019). Screening in contract design: evidence from the Aca health insurance exchanges. *American Economic Journal: Economic Policy*, *11*(2), 64-107.
10. Breyer, F., Bundorf K., and Pauly, M.V. 2012. "Health Care Spending Risk, Health Insurance, and Payment to Health Plans". In: Pauly, McGuire, T.G and Barros, P. (eds.). *The Handbook of Health Economic, Vol. 2*. Elsevier, pp 691-762.
11. Rice, N. and Smith, P.C. 2001. "Capitation and risk-adjustment in health care financing: an international progress report". *The Milbank Quarterly*, 79(1): 81-113.
12. Van de Ven, W. P. M. M. and Ellis, R. P. 2000. "Risk-adjustment in competitive health plan markets". In *Handbook of Health Economics* (Vol. 1, pp. 755-845). Elsevier.
13. Glazer, J. and McGuire, T.G. 2002. "Setting Health Plan Premiums to Ensure Efficient Quality in Health Care: Minimum Variance Optimal Risk-adjustment", *Journal of Public Economics*, 84: 153-173.
14. Newhouse, J. P. 2002. "Why is there a quality chasm?". *Health Affairs*, *21*(4), 13-25.
15. Van de Ven, W. P., Beck, K., Buchner, F., Chernichovsky, D., Gardiol, L., Holly, A. and Van de Voorde, C. 2003. "Risk-adjustment and risk selection on the sickness fund insurance market in five European countries". *Health Policy*, *65*(1), 75-98.
16. Van Barneveld, E. M., Lamers, L. M., Van Vliet, R. C., and Van de Ven, W. P. 2001a. "Risk-sharing as a supplement to imperfect capitation: a tradeoff between selection and efficiency". *Journal of Health Economics*, *20*(2), 147-168.

17. Van Barneveld, E. M., Van Vliet, R. C., and Van de Ven, W. P. 2001b. "Risk-sharing between competing health plans and sponsors". *Health Affairs*, *20*(3), 253-262.
18. Newhouse, J. P. 1996. "Reimbursing health plans and health providers: efficiency in production versus selection". *Journal of Economic Literature*, 1236-1263.
19. Schokkaert, E., Geert D., and Van De Voorde, C. 1998. "Risk-adjustment and the trade-off between efficiency and risk selection: an application of the theory of fair compensation." *Health Economics* 7(5): 465-480.
20. McGuire, T. and van Kleef, R. (eds). 2018. Risk-adjustment, Risk-sharing and Premium Regulation in Health Insurance Markets: Theory and Practice, Elsevier Publishing.
21. Van de Ven, W. P., Beck, K., Van de Voorde, C., Wasem, J., and Zmora, I. 2007. "Risk-adjustment and risk selection in Europe: 6 years later". *Health Policy*, *83*(2), 162-179.
22. Van de Ven, W. P., Beck, K., Buchner, F., Schokkaert, E., Schut, F. E., Shmueli, A., & Wasem, J. 2013. "Preconditions for efficiency and affordability in competitive healthcare markets: are they fulfilled in Belgium, Germany, Israel, the Netherlands and Switzerland?". *Health Policy*, *109*(3), 226-245.
23. Van de Ven, W. P., van Kleef, R. C., and van Vliet, R. C. 2015. "Risk selection threatens quality of care for certain patients: lessons from Europe's health insurance exchanges". *Health Affairs*, *34*(10), 1713-1720.
24. Van Veen, S. H. C. M., Van Kleef, R. C., Van de Ven, W. P. M. M., and Van Vliet, R. C. J. A. 2015. "Is there one measure-of-fit that fits all? A taxonomy and review of measures-of-fit for risk-equalization models". *Medical Care Research and Review*, *72*(2), 220-243.
25. Layton, T., Ellis, R., McGuire, T. and van Kleef, R. 2017. "Measuring Efficiency of Health Plan Payment Systems in Managed Competition Health Insurance Markets," *Journal of Health Economics*, (56): 237-255.
26. Geruso, M., & McGuire, T. G. (2016). Tradeoffs in the design of health plan payment systems: Fit, power and balance. *Journal of health economics*, 47, 1-19.
27. Brammli-Greenberg, S., Glazer, J., and Shmueli, A. 2018. "Regulated Competition and Health Plan Payment Under the National Health Insurance Law in Israel—The Unfinished Story". In *Risk-adjustment, Risk-sharing and Premium Regulation in Health Insurance Markets* (pp. 365-395). Academic Press.
28. Shmueli, A. and Nissan-Engelcin, E., 2013. "Local availability of physicians' services as a tool for implicit risk selection." *Social Science & Medicine*, 84, pp.53-60.