Data-Driven Incentive Alignment in Capitation Schemes

Mark Braverman Sylvain Chassang*†

Princeton University New York University

August 9, 2016

Abstract

This paper explores whether Big Data, taking the form of extensive but high dimensional records, can reduce the cost of adverse selection in government-run capitation schemes, such as Medicare Advantage, or school voucher programs. We argue that using data to improve the ex ante precision of capitation regressions is unlikely to be helpful. Even if types become essentially observable, the high dimensionality of covariates makes it infeasible to precisely estimate the cost of serving a given type. This gives an informed private operator scope to select types that are relatively cheap to serve. Instead, we argue that data can be used to align incentives by forming unbiased and non-manipulable ex post estimates of a private operator's gains from selection.

Keywords: adverse selection, big data, capitation, observable but not interpretable, health-care regulation, detail-free mechanism design, model selection.

1 Introduction

This paper explores the value of Big Data in reducing the cost of adverse selection in government-run capitation or voucher schemes, with a particular emphasis on healthcare

*Braverman acknowledges support from NSF Award CCF-1215990, NSF CAREER award CCF-1149888, a Turing Centenary Fellowship, and a Packard Fellowship in Science and Engineering. Chassang acknowledges support from the Alfred P. Sloan Foundation.

†We're grateful to Ben Brooks, Janet Currie, Mark Duggan, Kate Ho, Amanda Kowalski, Roger Myerson, Phil Reny, Dan Zeltzer as well as seminar participants at Boston University, Princeton, and the Becker-Friedman institute at the University of Chicago, for many helpful comments.

insurance.

Traditional capitation schemes pay private plans an estimate of expected public cost of service for each individual they enroll. Examples of of capitation schemes include Medicare Advantage, a program which lets US Medicare recipients switch to private health insurance plans, as well as school vouchers. Capitation payments can be conditioned on agreed upon user characteristics (then, they are said to be risk-adjusted). While capitation programs are a popular way to outsource government mandated services to the private sector, they are often plagued by adverse selection. Private service plans have strong incentives to select types that are cheaper to serve than their capitation payment, which increases the cost of serving the overall population. In the context of Medicare Advantage, Batata (2004) and Brown et al. (2014) report yearly overpayments in the thousands dollars for patients selected by private plans.

A natural strategy to reduce adverse selection is to build precise, risk-adjusted, ex ante capitation schemes, reimbursing private plans for the expected cost of taking care of the specific patients they select. This suggests that Big Data — i.e., the availability of high-dimensional patient records — which can be used to condition capitation payments on precise individual characteristics, may be of considerable help in reducing the effects of adverse selection. We take a different view and argue that under the correct Big Data limit, this naïve use of high-dimensional co-variates is likely to be of limited value. Instead, we suggest that data may be more successfully used to form unbiased ex post estimates of strategic selection by private plans. Correcting capitation formulas with these ex post estimates aligns the public and private plans' incentives.

Our model considers a single public plan seeking to outsource the provision of healthcare services to a single private plan.[1] The private plan may have a comparative advantage in treating certain types so that some selection of patients may be welfare enhancing. However the private plan also has incentives to select patients whose cost of care is mispriced. This

[1]In the case of Medicare Advantage, the private plan would correspond to a PPO or HMO.

creates a distinction between legitimate selection characteristics, which predict comparative advantage, and illegitimate selection characteristics, which predict costs but not comparative advantage. Efficient selection need only depend on legitimate selection characteristics.

Our modeling choices reflect both the opportunities and limitations presented by Big Data. We assume that high-dimensional records isomorphic to patients' types — i.e. sufficient statistics for patients' cost of care — are observable. However, we also recognize that the number of such possible types need not be small relative to the sample size of available cost data, thereby limiting their use for prediction. This leads us to study mechanism design at a joint limit where both the sample size and the number of relevant covariates are large.[2] At this Big Data limit, sufficient statistics of types are observable but not interpretable. This creates a trade-off when setting capitation rates: "sparse" cost estimates, conditioned on a few patient characteristics, have low standard errors but high bias; in contrast "rich" cost estimates, conditioned on an exhaustive set of patient characteristics, have low bias, but large standard errors.

The trade-off captured by our Big Data limit is reflected in the capitation schemes employed by Medicare Advantage, as well as in the risk-adjustment formula used to calculate transfers between plans under the Affordable Care Act (ACA). The Medicare Advantage risk adjustment model, rolled out in 2004, uses Hierarchical Condition Categories (HCC) (Pope et al., 2004). The HCCs are groups of conditions that can be inferred from the patient diagnosis data. The number of HCCs in the model varies between editions, but is generally under 100. They are used in conjunction with condition severity modifiers, and demographic factors to estimate individual patients' expected expenditures in the subsequent year. Thus the model falls under the "sparse capitation" type which we discuss below: there are relatively few categories, and thus a reasonably unbiased estimator can be formed for each category (Evans et al., 2011). In fact, the desire for "adequate sample sizes to permit accurate and

[2]This is the limit taken in the statistics literature concerned with Big Data. See Belloni et al. (2013, 2014) for recent examples in econometrics.

stable estimates of expenditures" has been a design principle for the risk adjustment scheme, and a factor in keeping the number of patient types in the model relatively low (Pope et al., 2004).

The model used for risk-adjustment transfers under the ACAs uses an adapted set of HCCs (since the ACA transfer model is a general-population model, while the Medicare Advantage model is primarily for the 65+ population) (Kautter et al., 2014). It uses 114 HCCs. As in the case of Medicare Advantage model, the need for statistical power to get ex-ante good estimates is one of the design principles limiting the number of categories used (for Medicare et al., 2016). An additional feature of the ACA risk adjustment scheme is that it is "budget-neutral" — one plan's gain under the scheme is another plan's loss, and there is no calibrating set held by the government. This introduces additional incentive issues which we address in Section 5.

Our first set of results considers traditional capitation schemes, which, as emphasized by Brown et al. (2014), seek to reimburse private plans for the expected cost of treating patients given ex ante observables. Sparse capitation schemes condition cost estimates on a small set of patient characteristics, while rich capitation schemes condition cost estimates on the full set of characteristics made available by Big Data. We show that such schemes induce efficient selection when capitation fees conditional on types are precisely estimated, or when the private plan is constrained to select only on the basis of legitimate characteristics. However, we show that these conditions fail under our Big Data limit. Indeed, cost-estimates conditional on types remain noisy even for large samples. Hence, even though types are observable, it is possible for the private plan to maintain an informational advantage which induces inefficient selection and increases the average cost of care.

In spite of these limitations, we are able to show that an appropriate ex post use of data can achieve efficient selection at no excess cost for the public plan whenever legitimate selection characteristics are common knowledge. Instead of including a large number of covariates to obtain a more precise capitation formula, we argue that it is sufficient to augment the baseline capitation formula (based on legitimate characteristics) with a single additional term measuring ex post selection by the private plan. This additional term takes the form of an appropriately weighted covariance between the distribution of types selected by the private plan, and the residuals from the basic capitation regression evaluated on out-of-sample costs. More concretely, it provides an unbiased estimate of the cost savings obtained by the private plan from selecting a non-representative sample of patients. This "strategic capitation scheme" induces efficient selection, and, importantly, does not give the public plan any incentive to bias its report of out-of-sample costs. This last property allows us to extend our approach to health exchanges for which out-of-sample cost realizations would be reported by competing healthcare plans (see extension in Section 5).

The basic idea behind strategic capitation can be extended to environments where legitimate selection characteristics are not common knowledge. In this case it is still possible to achieve a meaningful share of first-best efficiency by using generalized strategic capitation schemes that let private plans specify the characteristics they wish to select on. This flexibility comes at a cost related to the complexity of the class of models the private plan can use to select patients. We show that the performance guarantees of this scheme are essentially unimprovable by studying the exact direct mechanism design problem in specific environments.

The paper contributes to the theoretical literature on adverse selection in insurance markets.[3] Our work is particularly related to Glazer and McGuire (2000), who study optimal risk-adjustment in a Bayesian setting. They show that when selection is possible, optimal ex ante reimbursement schemes should deviate from simply reimbursing private plans the expected cost of taking care of patients. In particular, capitation schemes should adjust reimbursement rates to dull the effect of cream-skimming by private plans. We show how to induce efficient selection by using information about patient types and ex post cost data.

[3]See for instance Rothschild and Stiglitz (1976), Bisin and Gottardi (1999, 2006), Dubey and Geanakoplos (2002).

Our mechanism is closely related to that of Mezzetti (2004), which also uses noisy ex post information to provide accurate ex ante incentives. Also related is the work of Riordan and Sappington (1988) who show how to exploit noisy ex post signals to screen agents at no cost to the principal. As we clarify in greater detail later in the paper, our work differs for two main reasons. First, we are interested in prior-free mechanisms and do not make the identification assumptions required in Riordan and Sappington (1988). Second, ex post signals (here the public plan's hold-out cost data) need not be publicly observed and we must ensure that the relevant party has correct incentives for reporting. Third, unlike Mezzetti (2004), we require exact budget-balance.

Our work is motivated by a growing empirical literature which documents cream-skimming in health insurance and education markets, and studies the efficiency of various risk-adjustment schemes (Frank et al., 2000, Mello et al., 2003, Batata, 2004, Epple et al., 2004, Newhouse et al., 2012, Walters, 2012, Brown et al., 2014). Our analysis is largely inspired by Brown et al. (2014) which shows that increasing the number of covariates used in Medicare Advantage's capitation formulas has in fact led to an increase in the cost of adverse selection to the state.[4] We complement this result by showing that naïve uses of data are unlikely to resolve adverse selection, but suggest that progress can be made by using data to detect selection ex post.

The paper is structured as follows. Section 2 describes our framework, and in particular our approach to Big Data. Section 3 uses a simple example in which legitimate selection characteristics are common knowledge to delineate the mechanics of adverse selection under various capitation schemes. Section 4 generalizes the analysis to settings in which the private plan's comparative advantage is not common knowledge. Section 5 presents several extensions. We show how to adapt our approach to address adverse selection in markets with multiple private plans and no public plan. In addition, we briefly discuss how to address concerns of risk inflation, dynamic selection, and reduced quality provision by private plans.

[4]Newhouse et al. (2012) argues that the cost of adverse selection may be overstated.

Details are provided in Appendix A. Proofs are collected in Appendix B unless mentioned otherwise.

2 Framework

Our model seeks to capture three main features. The first is selection by private health-care plans, such as HMOs or PPOs, which we model as a reduced form cost for attracting different populations. Selection may be achieved through targeted advertisement and marketing (consistent with Starc (2014)), heterogenity in the quality of customer service during enrollment procedures, as well as targeted service bundles.

Second, public and private plans have heterogeneous comparative advantages in treating patients. Indeed, insurance plans serve a role beyond that of financial intermediaries. Plans play an important role in selecting, monitoring and generally resolving agency problems vis à vis doctors and hospitals, as well as encouraging preventive care and healthy habit formation. Data from Bundorf et al. (2012) provides evidence for such comparative advantage across different plans. In their sample, HMOs have a comparative advantage over PPOs in treating high risk patients. In our model, this creates a reason for both public and private plans to be active, and raises the question of efficient patient allocation.

Third, we seek to correctly capture the forces that make Big Data attractive but challenging: we assume that high dimensional records make patients' type observable, but that as a result, even with large samples of patients, it is not possible to form precise estimates of expected cost of treatment conditional on type (this concern for power is reflected in Pope et al. (2004), Evans et al. (2011), Kautter et al. (2014)). Types are observable but no interpretable.

The lead example for our work is Medicare Advantage, a program which lets US Medicare recipients switch to private insurance plans such as HMOs and PPOs. Medicare Advantage is a large and growing program. It covers a population of roughly 15 million, out of the roughly 50 million enrolled in Medicare, and its size was multiplied by three from 2005 to 2015. Selection by private plans is also an ongoing concern threatening the financial sustainability of the program (Batata, 2004, Brown et al., 2014).

2.1 Players, Actions, Payoffs

We study the relationship between a public health care plan p_0 responsible for the health expenses of a set $I = \{1, \cdots, N\}$ of patients and an independent private plan p_1.

Treatment costs. Each patient $i \in I$ has a type $\tau_i \in T \subset \mathbb{R}^n$ where the set of types T is potentially very large, but finite. Type τ is a sufficient statistic for the patient's cost of care. For any sample J of patients, we denote by $\mu_J \in \Delta(T)$ the sample distribution of types τ defined by $\mu_J(\tau) \equiv \frac{|J^\tau|}{|J|}$, where $J^\tau \equiv \{j \in J | \tau_j = \tau\}$, and $|J|$ denotes the cardinal of J.

Realized cost of care for a patient i of type τ, insured by plan p are denoted by $\widehat{c}_i(p) \geq 0$, and the corresponding sample distribution of costs conditional on τ and p is denoted by $c(\tau, p) \in \Delta(\mathbb{R}_+)$. Note that the sample distribution is itself uncertain. Treatment costs are exchangeable conditional on patient type τ and plan p.

We denote by $\mathbb{E}_c$ expectations under the realized sample distribution of costs c. Let $\kappa(\tau, p) \equiv \mathbb{E}_c[\widehat{c}|\tau, p]$ denote the expected realized cost of treatment for a patient of type τ by plan p, given sample distribution c, so that $\widehat{c}_i(p)$ can be written as

$$\widehat{c}_i(p) = \kappa(\tau_i, p) + e_{i,p}, \tag{1}$$

where $\mathbb{E}_c[e_{i,p}] = 0$.

To simplify welfare statements, we assume that the public and private plan share a common prior $\nu \in \Delta(\Delta(\mathbb{R}^{T \times \{p_0, p_1\}}))$ over costs c. Note that the capitation mechanisms we study do not rely on the common prior assumption. Our performance bounds remain valid in a non-common prior setting, if expectations are taken under the private plan's prior.

Selection. Private plan p_1 can choose an expected selection policy $\lambda : T \to [0, 1]$ at a cost $K(\lambda) \geq 0$. Consistent with observations in Starc (2014), this reduced-form cost of selection may be thought of as a cost of advertisement.[5] Realized selection $\Lambda \subset I$ is a mean preserving spread of intended selection λ defined by

$$\mathbf{1}_{i\in\Lambda} = \lambda(\tau_i) + \varphi_i$$

where error term $(\varphi_i)_{i\in I}$ has expectation equal to zero, and is independent of cost shocks $e_{i,p}$, but may be correlated across different types $\tau \in T$. For instance, recruitment ads may unexpectedly attract a population different from the targeted one.

Realized payoffs. Given a selection decision λ by private plan p_1, a realized selection Λ, and a transfer $\Pi \in \mathbb{R}$ from p_0 to p_1, the realized surpluses U_0 and U_1 accruing to the public and private plans are

$$U_0 = -\Pi + \sum_{i\in\Lambda} \widehat{c}_i(p_0) \quad \text{and} \quad U_1 = \Pi - \sum_{i\in\Lambda} \widehat{c}_i(p_1) - K(\lambda).$$

2.2 Data

We model explicitly the role that data plays in the contracting problem. In particular we formalize a "Big Data limit" which captures the idea that although types are observable, when the type-space is large, the public plan may still have very imprecise estimates of expected treatment costs conditional on types. A consequence illustrated in Section 3 is that imprecise additional signals may give the private plan a significant advantage in selecting patients.

[5]Under a more standard model of selection along the lines of Rothschild and Stiglitz (1976), the private plan would screen patients through a menu of discounts and benefits specifically appealing to desirable types.

Samples. Both plans p_0 and p_1 observe a public dataset of types and cost realizations $D_0 = \{(i, \tau_i, \widehat{c}_i(p_0))|i \in D_0\}$ for plan p_0, where $i \in D_0$ denotes a patient i whose record is included in D_0. In addition, we denote by $D_0^\tau = \{(i, \tau_i, \widehat{c}_i(p_0))|\tau_i = \tau, i \in D_0\}$ the cost data relating to patients of type τ. We assume that for every $\tau \in T$, the set D_0^τ is non-empty, which implies $|T| \leq |D_0|$: the sample size of dataset D_0 is at least as large as the type space.

Plan p_1 privately observes a dataset $D_1 = \{(i, x_i, \widehat{c}_i(p_1))|i \in D_1\}$ reporting both her own costs, and side-signals x_i for a sample of patients $i \in D_1$. Side signal x_i captures other signals beyond cost realizations that the plan may be able to use in order to select patients.

Finally, we assume that plan p_0 has access to a hold-out sample $H = \{(i, \tau_i, \widehat{c}_i(p_0))|i \in H\}$ of her own costs, independent of data D_1 conditional on the realization of cost distribution $(c(\tau, p))_{\tau \in I, p \in \{p_0, p_1\}}$. Hold-out sample H may consist of ex post cost realizations for the current set of patients enrolled by the public plan. Alternatively, H may correspond to past cost data, securely encrypted, and verifiably released only after patient selection has occurred.[6] Contracts will be allowed to depend on hold-out sample H, but we will take seriously the public plan's incentive to reveal correct information.[7] Access to such hold-out sample data is essential. It allows the public plan to obtain estimates of her own costs whose errors are uncorrelated to the private plan's private information.

Big Data. Our model of Big Data consists of two assumptions (recall that $\mu_I \in \Delta(T)$ is the sample distribution of types τ in the patient population):

(i) types $\tau \in T$ are publicly observable;

(ii) sample data D_0, type space T and sample I grow large together, so that

$$\limsup_{|D_0| \to \infty} \frac{|I|}{|D_0|} < \infty \quad \text{and} \quad \liminf_{|D_0| \to \infty} \mathbb{E}_{\mu_I} \left[\frac{1}{|D_0^\tau|} \right] > 0.$$

[6]For instance, an encrypted version of the data can be released before selection occurs, with a decryption key publicized after patient enrollment has occurred.

[7]Specifically, we will address the public plan's incentives to bias its records in order to reduce payments to the private plan. For instance the public plan could down-code interventions happening to its own patients

Points (i) and (ii) summarize what we think are the opportunities and limitations of Big Data. On the one hand, high dimensional records make types observable (i). On the other hand, even though the aggregate sample size D_0 is large, the state space T is not small compared to D_0. Under sample measure μ_I, the size $|D_0^\tau|$ of sufficiently many subgroups D_0^τ remains bounded above, which implies that public plan p_0's estimates of costs on the basis of data D_0 necessarily remain noisy. We note that for the results in this paper to hold, the first condition in (ii) can be replaced with the weaker $\lim_{|I|\to\infty} |D_0| = \infty$, even though we believe the condition as stated to be realistic.

Note that since type space T is changing, the limit described above considers sequences of models. It should be treated as a stylized approximation capturing the fact that in the existing data, the number of a priori relevant characteristics (or columns) is not small compared to the number of data points (or rows). Throughout the paper, we provide bounds that depend explicitly on $\frac{|I|}{|D_0|}$ and $\mathbb{E}_{\mu_I}\left[\frac{1}{|D_0^\tau|}\right]$.

2.3 Contracts, Equilibrium and Welfare

Contracts. For any set of patients $J \subset I$, let $\tau_J \equiv (\tau_i)_{i\in J}$ and $\widehat{c}_J(p) \equiv (\widehat{c}_i(p))_{i\in J}$ denote profiles of types and costs. We denote by $H_R = \{(i, \tau_i, \widehat{c}_i^R(p_0)), i \in H\}$ the hold-out data reported ex post by p_0. We emphasize that these are reports of privately observed costs, and that the public plan must be given incentives to report truthfully. A capitation contract between the public and private plan is a mapping $\Pi(D_0, \Lambda, \tau_I, H_R) \in \mathbb{R}$, specifying the aggregate payments received by private plan p_1 as a function of public data D_0, realized selection Λ, the distribution of types τ_I in patient population I, and reported hold-out sample data H_R.

Equilibrium. We denote by β the public plan's strategy, mapping hold-out data H to reported hold-out data H_R. Given a capitation contract Π, a selection strategy λ, and a reporting strategy β, the public and private plans obtain expected payoffs,

$$\mathbb{E}_\nu U_0 = \mathbb{E}_\nu \left[-\Pi + \sum_{i\in\Lambda} \widehat{c}_i(p_0) \middle| \lambda, \beta \right],$$
$$\mathbb{E}_\nu U_1 = \mathbb{E}_\nu \left[\Pi - \sum_{i\in\Lambda} \widehat{c}_i(p_1) \middle| \lambda, \beta \right] - K(\lambda).$$

Given a contract Π, abstractly denoting by $\mathcal{I}_0$ and $\mathcal{I}_1$ the information available to plans p_0 and p_1, a strategy profile (β, λ) is in equilibrium if and only if β and λ respectively solve

$$\max_\beta \mathbb{E}_\nu[-\Pi \middle| \mathcal{I}_0, \beta, \lambda] \quad \text{and} \quad \max_\lambda \mathbb{E}_\nu \left[\Pi - \sum_{i\in\Lambda} \widehat{c}_i(p_1) \middle| \mathcal{I}_1, \beta, \lambda \right] - K(\lambda).$$

We denote by $\beta^*(H) \equiv H$ the truthful reporting strategy. We break indifferences in favor of truthful reporting, i.e. we assume that plan p_0 sends truthful reports whenever it is an optimal strategy, reflecting small costs in misreporting.

Design objectives. Conditional on selection rule λ and expected costs κ, surplus takes the form

$$S(\lambda) = -K(\lambda) + \sum_{i\in I} \lambda(\tau_i) \left[\kappa(p_0, \tau_i) - \kappa(p_1, \tau_i) \right].$$

We seek contracts Π such that for all priors ν, data D_0, D_1, and all equilibria (λ, β):

$$\mathbb{E}_\nu[S|\lambda] = \mathbb{E}_{D_0,D_1\sim\nu} \left[\max_\lambda \mathbb{E}_\nu[S|\lambda, D_0, D_1] \right] - o(|I|) \tag{2}$$
$$\mathbb{E}_\nu \left[U_0 \middle| \lambda, \beta, D_0 \right] \geq -o(|I|) \tag{3}$$
$$\mathbb{E}_\nu \left[U_1 \middle| \lambda, \beta, D_1 \right] \geq 0. \tag{4}$$

In other terms, we seek ex post budget-balanced prior-free mechanisms that: maximize efficiency given available information up to a term negligible compared to the size $|I|$ of the patient population; satisfy at least approximate interim individual rationality for both plans. We highlight once again that the mechanisms we propose to attain these objectives do not exploit the common prior assumption, and would satisfy the same properties in a non-common prior setting, with expectations evaluated under the private plan's prior.[8]

3 An Example

To fix ideas, we delineate our main points using a simple instantiation of the model introduced in Section 2.

Legitimate and illegitimate selection. We assume in this example that there exists a common knowledge partition E of type space T, with typical element $\eta \in E$ (so that $\eta \subset T$ is a subset of T, e.g. the set of patient sharing a commom medical condition) such that treatment costs can be decomposed as

$$\widehat{c}_i(p) = \kappa(\eta_i, p) + e_{i,\tau_i} \tag{5}$$

where terms $e_{i,\tau}$ have mean zero conditional on η, and are distributed according to a log-normal distribution:

$$e_{i,\tau} = \underline{\kappa} \left[\exp\left(\varepsilon_\tau + \varepsilon_i - 1\right) - 1\right]$$

with ε_τ and ε_i independent standard normal distributions $\mathcal{N}(0,1)$, and $\underline{\kappa} \in \left(0, \min_{\eta \in E, p \in \{p_0, p_1\}} \kappa(\eta, p)\right)$. By construction, $\mathbb{E}_\nu[e_{i,\tau}] = 0$ and $\widehat{c}_i(p) \geq 0$.[9]

Cost decomposition (5) is a special case of decomposition (1) in which the comparative advantages of plans p_0 and p_1, described by $\kappa(\eta, \cdot)$, depend only on characteristics $\eta \in E$. We think of E as a small set compared to T, so that it is possible for each plan to form

[8]For recent work emphasizing prior-free approaches to mechanism design, see Segal (2003), Bergemann and Schlag (2008), Hartline and Roughgarden (2008), Chassang (2013), Carroll (2013), Madarász and Prat (2014), Brooks (2014), Antic (2014).

[9]Throughout this example, we use the fact that a log-normal distribution $\ln \mathcal{N}(\mu, \sigma^2)$ has expectation $\exp\left(\mu + \frac{1}{2}\sigma^2\right)$.

accurate estimates of its costs conditional on $\eta \in E$. For simplicity, we assume that the costs of the public plan $\kappa(\cdot, p_0)$ are known by both plans, and that private plan p_1 knows its own costs $\kappa(\cdot, p_1)$. Error term $e_{i,\tau}$ captures residuals in cost estimates that depend both on idiosyncratic shocks ε_i, and type-level shocks ε_τ.

We assume in this example that the private plan is able to perfectly select the realized set Λ of patients it treats at no cost. That is, for all $\lambda \in [0,1]^T$, $K(\lambda) = 0$. An immediate implication of costless selection and cost decomposition (5) is that surplus maximizing selection rules need only depend on characteristics $\eta \in E$.

Remark 1. *First-best surplus, defined by* $S_{\max} \equiv \max_\lambda \mathbb{E}_c \left[\sum_{i \in \Lambda} \widehat{c}_i(p_0) - \widehat{c}_i(p_1) \middle| \lambda \right]$ *is attained by a selection policy* λ^* *that is measurable with respect to partition* E*:* $\lambda^*(\eta) = \mathbf{1}_{\kappa(\eta,p_0) > \kappa(\eta,p_1)}$.

Accordingly, a selection rule is said to be *legitimate* if and only if it is measurable with respect to E. Selection rules that are not measurable with respect to type-space partition E depend on features of types τ that do not matter for efficiency. They are referred to as *illegitimate*. We denote by $\mathcal{M}(E)$ the set of selection rules measurable with respect to E.

Private information. For every $\tau \in T$, the private plan's data D_1 lets it observe a signal $x_\tau = \varepsilon_\tau + \varepsilon_x$ with ε_x an independent error term distributed according to a standard normal $\mathcal{N}(0,1)$. Given that plan p_1 knows her expected costs $\kappa(\eta, p_1)$ this is equivalent to observing a single additional realization of her own costs $\widehat{c}_i(p_1)$ for each type $\tau_i \in T$.

Bayesian updating. The information structure defined above leads to tractable updated beliefs. Observing data D_0^τ is equivalent to observing signals $x_i = \varepsilon_\tau + \varepsilon_i$ for $i \in D_0^\tau$. Hence the public and private plan's beliefs over random cost parameter ε_τ follow normal distributions $\left(\mathcal{N}(\chi_{p,\tau}, \rho_{p,\tau}^{-1})\right)_{p \in \{p_0, p_1\}}$ where mean χ and precision ρ satisfy

$$\chi_{p,\tau} = \frac{\mathbf{1}_{p=p_1} x_\tau + \sum_{i \in D_0^\tau} x_i}{1 + \mathbf{1}_{p=p_1} + |D_0^\tau|} \quad \text{and} \quad \rho_{p,\tau} = 1 + \mathbf{1}_{p=p_1} + |D_0^\tau|. \tag{6}$$

This implies conditional estimates of residual costs

$$\mathbb{E}_\nu[e_{i,\tau}|D_0^\tau, p] = \underline{\kappa}\left[\exp\left(\chi_{p,\tau} - \frac{1}{2\left(|D_0^\tau| + 1 + \mathbf{1}_{p=p_0}\right)}\right) - 1\right].$$

Note that conditional on sample size $|D_0^\tau|$, precision $\rho_{p,\tau}$ is deterministic, while mean $\chi_{p,\tau}$ has an ex ante distribution $\mathcal{N}\left(0, \frac{\left(\mathbf{1}_{p=p_1}+|D_0^\tau|\right)^2+\mathbf{1}_{p=p_1}+|D_0^\tau|}{\left(1+\mathbf{1}_{p=p_1}+|D_0^\tau|\right)^2}\right)$. Term $\mathbf{1}_{p=p_1}$ corresponds to the informational advantage private plan p_1 derives from observing an additional cost realization.

3.1 Why Ex Ante Capitation Schemes Fail

We begin by illustrating the limits of natural transfer schemes that attempt to align incentives through fixed capitation rates. Since payments are specified ex ante, such mechanisms remove concerns that the public plan may misreport its hold-out costs to reduce payments. We show that under restrictive strategic environments, these schemes can indeed attain efficiency and satisfy both plans' individual rationality constraints. However, whenever plan p_1 can engage in illegitimate selection, these ex ante schemes are inefficient and generate large losses for public plan p_0.

We consider sparse and rich capitation contracts that differ in the sophistication of the regressions used to predict treatment costs. Transfers take one of the following forms:

$$\Pi^{\mathsf{sparse}}(\Lambda, \tau_I) = \sum_{i\in\Lambda} \mathbb{E}_\nu[c(\tau_i, p_0)|\eta_i, D_0] = \sum_{i\in\Lambda} \kappa(\eta_i, p_0) \tag{7}$$

$$\Pi^{\mathsf{rich}}(\Lambda, \tau_I) = \sum_{i\in\Lambda} \mathbb{E}_\nu[c(\tau_i, p_0)|\tau_i, D_0] = \sum_{i\in\Lambda} \kappa(\eta_i, p_0) + \mathbb{E}_\nu[e_{i,\tau_i}|\tau_i, D_0]. \tag{8}$$

In both schemes the private plan is paid the public plan's expected cost of treating selected patients, conditional on some set of ex ante observables. Note that since the private plan is the residual claimant of costs, it has incentives to provide required care as efficiently as possible. Sparse capitation estimates patients' costs conditional on legitimate characteristics η alone. Rich capitation estimates patients' costs conditional on the full set of observables

τ — i.e. it exploits Big Data to form targeted estimates. We now show that neither scheme resolves the problem of adverse selection at the Big Data limit.

Proposition 1 (sparse capitation)**.** *Consider capitation scheme* Π^{sparse}*.*

(i) Assume that the private plan is constrained to use legitimate selection, i.e. selection strategies λ *must be measurable with respect to* E*. Efficient selection and truthful reporting* (λ^*, β^*) *is the unique equilibrium.*

(ii) Assume that the private plan is not constrained to use legitimate selection rules. If $\mathbb{E}_{\mu_I}(|\kappa(\eta, p_0) - \kappa(\eta, p_1)|) > 0$*, then there exists* $h > 0$ *such that for all sample sizes* $|D_0|$*, scheme* Π^{sparse} *induces an efficiency loss*

$$\mathbb{E}_\nu \left[S_{\max} - S^{\mathsf{sparse}}\right] \geq h|I|.$$

If allocation does not matter for efficiency, i.e. $\forall \eta \in E$*,* $\kappa(\eta, p_0) = \kappa(\eta, p_1)$*, the public plan makes expected losses*

$$\mathbb{E}_\nu[U_0|\Pi^{\mathsf{sparse}}] \leq -h|I|.$$

If plan p_1 uses only legitimate selection strategies, sparse capitation induces efficient selection. Indeed the expected benefit that plan p_1 obtains from selecting in a patient with characteristic η is equal to $\kappa(\eta, p_0) - \kappa(\eta, p_1)$. This gives the private plan incentives to engage in efficient selection.

However, the profit a private plan p_1 expects from selecting a patient of type $\tau \in \eta$ is in fact $\kappa(\eta, p_0) - \kappa(\eta, p_1) - \mathbb{E}_\nu[e_{i,\tau}|D_0, x]$. Whenever the private plan can select on the basis of non-legitimate characteristics τ, term $\mathbb{E}_\nu[e_{i,\tau}|D_0, x]$ will induce deviations from efficient selection to avoid under-reimbursed patients and recruit over-reimbursed patients. This inefficiency arises because of bias in cost estimates, and does not vanish as the data gets large. Indeed, as we have observed, $\mathbb{E}_\nu[e_{i,\tau}|D_0,x] = \underline{\kappa}\left[\exp\left(\chi_\tau - \frac{1}{2}\frac{1}{|D_0^\tau|+2}\right) - 1\right]$ with χ_τ following a Gaussian distribution $\mathcal{N}\left(0, \frac{(|D_0^\tau|+1)^2+|D_0^\tau|+1}{(|D_0^\tau|+2)^2}\right)$. Hence perturbation $\mathbb{E}_\nu[e_{i,\tau}|D_0,x]$ and the inefficiency loss it induces do not vanish as sample size $|D_0|$ grows large.

Since inefficiencies in sparse capitation schemes are driven by biased cost estimates, rich capitation schemes Π^{rich}, which condition capitation rates on the full set of observables τ emerge naturally as a candidate solution. The following holds.

Proposition 2 (rich capitation)**.** *There exists continuous, strictly increasing, functions $\underline{h}$ and $\overline{h}$ satisfying $\underline{h}(0) = \overline{h}(0) = 0$, such that for all sample size distributions $(|D_0^\tau|)_{\tau\in T}$:*

(i) efficiency loss $S_{\max} - S^{\mathsf{rich}}$ satisfies

$$\mathbb{E}_\nu\left[S_{\max} - S^{\mathsf{rich}}\right] \le \underline{\kappa}\,|I|\,\overline{h}\left(\mathbb{E}_{\mu_I}\frac{1}{|D_0^\tau|}\right); \tag{9}$$

(ii) there exist mappings $\kappa(\cdot,p_0)$, $\kappa(\cdot,p_1)$ such that

$$\mathbb{E}_\nu\left[S_{\max} - S^{\mathsf{rich}}\right] \ge \underline{\kappa}\,|I|\,\underline{h}\left(\mathbb{E}_{\mu_I}\frac{1}{|D_0^\tau|}\right). \tag{10}$$

If $\kappa(\eta,p_0) = \kappa(\eta,p_1)$ for all η, the public plan makes expected losses

$$\mathbb{E}_\nu[U_0|D_0,\Pi^{\mathsf{rich}}] \le -\underline{\kappa}\,|I|\,\underline{h}\left(\mathbb{E}_{\mu_I}\frac{1}{|D_0^\tau|}\right).$$

While sparse capitation schemes do not achieve efficiency, regardless of data D_0, rich capitation schemes may achieve efficiency provided that $\mathbb{E}_{\mu_I}\frac{1}{|D_0^\tau|}$ becomes arbitrarily small, i.e. for almost every type τ, subsample D_0^τ becomes arbitrarily large. This is ruled out by definition at the Big Data limit. As a result, cost estimates $\mathbb{E}_\nu[e_{\tau,i}|D_0]$ remain imprecise for a non-vanishing mass of types τ (under patient sample measure μ_I) and signals $(x_\tau)_{\tau\in T}$ make it possible for private plan p_1 to profit from selecting mispriced types.

3.2 Strategic Capitation

We now describe a capitation scheme that correctly takes care of incentives for strategic selection by p_1 and strategic reporting by p_0. Payments can be expressed as

$$\Pi^{\mathsf{strat}}(\Lambda, \tau_I, H_R) \equiv \sum_{i \in \Lambda} \pi(\eta_i) + \Delta\pi(\eta_i, H_R), \tag{11}$$

where $\pi(\eta) \equiv \kappa(\eta, p_0)$ is the baseline capitation rate conditional on legitimate characteristics used in sparse capitation, and $\Delta\pi(\eta, H_R)$ is a correction dependent on reported hold-out data H^R and selected sample Λ taking the form:

$$\Delta\pi(\eta_i, H_R, \Lambda) \equiv \mathrm{cov}_I(s_i, r_i | \eta_i = \eta) = \frac{1}{|I^\eta|} \sum_{i \in I^\eta} s_i r_i, \quad \text{where}$$

- $s_i \equiv \frac{\mu_\Lambda(\tau_i|\eta_i)}{\mu_I(\tau_i|\eta_i)} - 1$ is a measure of selected sample Λ's deviation from legitimate selection;[10]

- $r_i \equiv \frac{1}{|H_R^{\tau_i}|} \sum_{j \in H_R^{\tau_i}} \left[\widehat{c}_j^R(p_0) - \kappa(\eta, p_0)\right]$ is the average residual of costs for type τ_i in the reported hold-out sample $H_R^{\tau_i} \equiv \{(j, \tau_j, \widehat{c}_j^R) | j \in H_R, \tau_j = \tau_i\}$.

Strategic capitation satisfies the following key properties

$$\forall\lambda, \quad \mathbb{E}_\nu[\Delta\pi(\eta_i, H_R, \Lambda)|D_0, D_1, \beta^*, \lambda] = \mathbb{E}_\nu[(\lambda(\tau|\eta) - \mu_I(\tau|\eta))\mathbb{E}_\nu[e_{\tau,i}|D_1, D_0]] \tag{12}$$

$$\forall\lambda \in \mathcal{M}(E), \forall\beta, \quad \mathbb{E}_\nu[\Delta\pi(\eta_i, H_R, \Lambda)|D_0, D_1, \beta, \lambda] = 0. \tag{13}$$

Condition (12) implies that under truthful reporting β^*, the adjustment performed by strategic capitation is an unbiased estimate of the excess profits plan p_1 may have obtained through illegitimate selection (the adjustment is negative if private plan p_1 overselects types that are comparatively cheaper to treat). This noisy ex post estimate provides an accurate ex ante

[10]Recall that for any sample J, $\mu_J(\tau|\eta) \equiv \frac{|J^\tau|}{|J^\eta|}$ denotes the distribution of types τ conditional on characteristic $\eta \subset T$ in sample J.

correction and dissuades inefficient selection. Condition (13) ensures that regardless of the public plan's reporting strategy β, the private plan can guarantee herself expected capitation payments $\pi(\eta) = \kappa(\eta, p_0)$, provided it uses a legitimate selection strategy $\lambda \in \mathcal{M}(E)$. [11]

Proposition 3. *Strategic capitation contract Π^{strat} induces a unique equilibrium (λ^*, β^*) in which private plan p_1 selects patients efficiently, and the public plan p_0 truthfully reports hold-out sample H. Both plans get positive expected payoffs: $\mathbb{E}_\nu[U_0|D_0, D_1, \lambda^*, \beta^*] \geq 0$ and $\mathbb{E}_\nu[U_1|D_0, D_1, \lambda^*, \beta^*] \geq 0$.*

Note that the observability of types τ is needed to assemble the correct cost residuals from the hold-out data, as well as to measure the private plan's deviation from legitimate selection. The hold-out sample is needed to ensure that residuals r_i are uncorrelated to plan p_1's information.

3.3 Alternative Mechanisms

To clarify the economic forces at work in our environment it is useful to delineate the mechanics of other relevant mechanisms.

Mechanisms from the literature. Other work has emphasized the value of ex post noisy signals in environments with quasi-linear preferences. Riordan and Sappington (1988) show that it is possible to efficiently regulate a monopoly with unknown costs by exploiting public signals correlated to the monopoly's type. Using a construction related to that of Cremer and McLean (1988), they show how to extract all the surplus by offering the monopoly appropriately chosen screening contracts. Strategic capitation also exploits the fact that noisy ex post signals (here, hold-out cost realizations) can be used to construct accurate ex ante incentives, but our environment differs in key ways. First, signals are not public, and we need to take care of the public plan's incentives to reveal its own cost. Second,

[11] This point plays a key role when studying incentives for truthful revelation in exchanges.

the identification condition at the heart of Riordan and Sappington (1988) is not satisfied: neither the distribution of the public plan's cost, nor the private plan's beliefs thereover, are sufficient statistic of the private plans' costs.

Mezzetti (2004) shows that it is possible to obtain efficiency in common value environments using ex post reports of the players' realized payoffs. In our application the mechanism proposed by Mezzetti (2004) would proceed by making the private plan a negative ex post transfer equal to the public plan's realized cost, and making the private plan a positive ex ante transfer to cover expected costs. This mechanism does not satisfy budget balance and relies on priors over the realized allocation to set ex ante transfers.

The differences between our environment and that of Mezzetti (2004) help clarify the role played by the Big Data assumption, i.e. the assumption that types are observable but not interpretable. We obtain budget balance by: forming a measure of the private plan's deviation from legitimate selection; interacting this measure with an unbiased estimate of the public plan's counterfactual costs. This ensures that in equilibrium, neither the private nor the public plan can affect their expected payoffs by deviating from legitimate selection and truthful reporting. The observability of types is used to compute the private plan's deviation from legitimate selection, as well as correctly reweight the distribution of types in the hold-out sample H to obtain estimates of counterfactual costs in the sample Λ of patients selected by the private plan.[12]

Plausible alternative mechanisms. A key step in strategic capitation is to use hold-out data to form estimates of counterfactual costs for the public plan. The assumption that types are observable is needed to reweight the distribution of types in the hold-out sample to match that of the selected sample. There may be other ways to form an unbiased estimate of counterfactuals. For instance, if it were possible to assign patients selected by the private plan back to the public plan with a fixed uniform probability, one could form an estimate

[12]The distribution of types in H and Λ should typically be different. For instance, the hold-out sample may consist of types treated by the public plan and rejected by the private plan.

of counterfactual costs without observing types. Beyond feasibility issues, a difficulty with this approach is that it does not take care of the public plan's incentives to bias its own cost reports.

Strategic capitation dissuades illegitimate selection by forming unbiased estimates of the private plan's excess profits. An alternative way to dissuade illegitimate selection is to impose sufficiently large penalties, say proportional to $\left|\frac{\mu_\Lambda(\tau_i|\eta_i)}{\mu_I(\tau_i|\eta_i)} - 1\right|$, when the sample selected by the private plan deviates from legitimate selection. This scheme requires the observability of types but does not require the availability of a hold-out sample. It induces efficient legitimate selection whenever the private plan can select patients precisely and at no cost. However this scheme carries an efficiency loss if it is costly to ensure that realized selection Λ is legitimate. Strategic capitation avoids the issue by using hold-out data to form an unbiased estimate of the profits from selection.

4 General Analysis

The strategic capitation scheme presented in Section 3 relies on strong assumptions. Chief among those, cost decomposition (5) ensures that the surplus maximizing policy depends on a small number of commonly known characteristics $\eta \in E$. This is not realistic: a private plan's comparative advantage is likely to be her private information, and it need not be the case that the optimal selection policy is measurable with respect to a small set of characteristics. Furthermore, private plans may be able to innovate and develop comparative advantages along new dimensions. Finally, in practice, the public plan's expected cost of treatment conditional on a characteristic η will have to be estimated from data. This creates additional room for selection by the private plan. This section extend strategic capitation to such environments.

We assume for simplicity that realized costs are bounded, i.e. that there exists $c_{\max}$ such that $\widehat{c}_i(p) \in [0, c_{\max}]$. Recall that $\kappa(\tau, p) = \mathbb{E}_c[\widehat{c}|\tau, p]$ denotes expected costs of treatment given τ, which yields decomposition $\widehat{c}_i(p) = \kappa(\tau_i, p) + e_i$, where $\mathbb{E}_\nu[e_i|\tau, p] = 0$. By construction, it must be that $e_i \in [-c_{\max}, c_{\max}]$. Finally, let

$$S(\lambda|D_0, D_1) \equiv \mathbb{E}_\nu \left[\sum_{i\in I} \lambda(\tau_i) \left[\kappa(p_0, \tau_i) - \kappa(p_1, \tau_i)\right] \Big| D_0, D_1 \right] - K(\lambda)$$

$$S_{E|D_0,D_1} \equiv \max_{\lambda\in\mathcal{M}(E)} S(\lambda|D_0, D_1)$$

respectively denote the surplus achieved by selection rule λ, and the maximum surplus achievable using selection rules measurable with respect to partition E.

4.1 Generalized Strategic Capitation

For any collection $\mathcal{E}$ of partitions $E \in \mathcal{E}$, our goal is to approach the maximum achievable efficiency $S_{E|D_0,D_1}$ with respect to partitions $E \in \mathcal{E}$. One difficulty is that the public plan's expected cost of treatment conditional on a characteristic $\eta \in E \in \mathcal{E}$ is no longer common knowledge. Instead, it must now be estimated from data. We define the generalized strategic capitation scheme $G^{\mathsf{strat}}_{\mathcal{E}}$ as follows:

1. data D_0 is shared with plan p_1;

2. plan p_1 picks a partition $E \in \mathcal{E}$ according to which it will be allowed to select patients; we continue to refer as characteristics $\eta \in E$ as legitimate selection characteristics;

3. plan p_1 is rewarded using the strategic capitation scheme Π^{strat} defined by

$$\Pi^{\mathsf{strat}}(\Lambda, \tau_I, H_R) \equiv \sum_{i\in\Lambda} \pi(\eta_i) + \Delta\pi(\eta_i, H_R)$$

   where $\pi(\eta) = \widehat{\kappa}(\eta, p_0) \equiv \sum_{\tau\in\eta} \mu_I(\tau|\eta) \frac{1}{|D_0^\tau|} \sum_{i\in D_0^\tau} \widehat{c}_i(p_0)$ is the sample estimate $\widehat{\kappa}(\eta, p_0)$ of the public plan's expected treatment costs conditional on characteristic $\eta \in E$. As in Section 3, $\Delta\pi(\eta, H_R, \Lambda)$ takes the form:

$$\Delta\pi(\eta_i, H_R, \Lambda) \equiv \mathrm{cov}_I(s_i, r_i|\eta_i = \eta) = \frac{1}{|I^\eta|}\sum_{i\in I^\eta} s_i r_i,$$

with

$$s_i \equiv \frac{\mu_\Lambda(\tau_i|\eta_i)}{\mu_I(\tau_i|\eta_i)} - 1 \quad \text{and} \quad r_i \equiv \frac{1}{|H_R^{\tau_i}|}\sum_{j\in H_R^{\tau_i}} \left[\widehat{c}_j^R(p_0) - \widehat{\kappa}(\eta, p_0)\right].$$

An equilibrium of mechanism $G_{\mathcal{E}}^{\mathsf{strat}}$ is a triplet (E, λ, β) where $E \in \mathcal{E}$ is p_1's choice of characteristics it is allowed to select on.

Mechanism $G_{\mathcal{E}}^{\mathsf{strat}}$ expands on strategic capitation by letting the private plan specify the set of characteristics it wishes to select on. As we show below, this additional degree of freedom results in unavoidable losses related to the complexity of the class of models $\mathcal{E}$ the private plan is allowed to pick from. These losses are related to penalties encountered in the model selection literature (Vapnik, 1998, Massart and Picard, 2007), and indeed one can think of our problem as one of delegated model selection.

Definition 1. *For any class of partitions $\mathcal{E}$ and error random variables $e = (e_i)_{i\in D_0}$, let $\Psi(\mathcal{E}, e)$ denote the random variable*

$$\Psi(\mathcal{E}, e) \equiv \max_{E\in\mathcal{E}} \left(\sum_{\eta\in E} |I^\eta| \left[\sum_{\tau\in\eta} \mu_I(\tau|\eta)\frac{1}{|D_0^\tau|}\sum_{i\in D_0^\tau} e_i\right]^+\right). \tag{14}$$

Variable $\Psi(\mathcal{E}, e)$ is an upper-bound to the gains a perfectly informed private plan could obtain from selecting the partition E that lets her optimally target over-reimbursed types. The scope for selection comes from the fact that generalized capitation uses sample averages $\widehat{\kappa}(\eta, p_0)$ to estimate the public plan's cost of service $\mathbb{E}_\nu[c_i(p_0)|\eta, c]$ conditional on legitimate characteristics.

Generalized capitation extends the performance bounds described in Proposition 3 up to a penalty of order $\mathbb{E}_\nu[\Psi(\mathcal{E}, e)]$.

Proposition 4 (efficiency bounds)**.** *Consider a collection of $\mathcal{E}$ of partitions. In any equilibrium (E, λ, β) of mechanism $G_{\mathcal{E}}^{\mathsf{strat}}$ we have that*

$$S(\lambda) \geq \mathbb{E}_\nu \left[\max_{E \in \mathcal{E}} S_{E|D_0, D_1} \right] - 2\mathbb{E}_\nu \left[\Psi(\mathcal{E}, e) \right]; \tag{15}$$

$$\mathbb{E}_\nu \left[-\Pi + \sum_{i \in \Lambda} \widehat{c}_i(p_0) \middle| D_0 \right] \geq -\mathbb{E}_\nu \left[\Psi(\mathcal{E}, e) \right]; \tag{16}$$

$$\mathbb{E}_\nu \left[\Pi - \sum_{i \in \Lambda} \widehat{c}_i(p_1) \middle| D_0, D_1 \right] \geq 0. \tag{17}$$

We do not endogenize the choice of the class of models $\mathcal{E}$. Still, if institutions are designed at a sufficiently ex ante period — specifically before data D_0 is realized — penalties $\Psi(\mathcal{E}, e)$ can be used to do so. The idea would be to let the private plan submit a class of models $\mathcal{E}$ ex ante that it will be able to pick from at the interim stage, and charge her complexity penalty $\mathbb{E}_\nu[\Psi(\mathcal{E}, e)]$. If data D_0 is renewed over time, the private plan may also be allowed to submit preferences over the class of models $\mathcal{E}$ to be used in the future.

Note that $\mathbb{E}_\nu[\Psi(\mathcal{E}, e)]$ depends on prior ν through error term e. The next lemma provides prior-free bounds for $\mathbb{E}_\nu \left[\Psi(\mathcal{E}, e) \right]$. Denote by $\alpha \equiv \mathbb{E}_{\mu_I} \left[\frac{|I^\tau|}{|D_0^\tau|} \frac{|D_0|}{|I|} \right] \geq 1$ the average representativeness of data D_0 for patients in I.[13] Let $M \equiv \sum_{E \in \mathcal{E}} \left(2^{|E|} - 1 \right)$.

Lemma 1 (selection bounds)**.** *(i) Let $(e_i')_{i \in I}$ denote i.i.d. Rademacher random variables uniformly distributed over $\{-c_{max}, c_{max}\}$. For any class $\mathcal{E}$ and any centered error terms $(e_i)_{i \in I}$ arbitrarily distributed over $[-c_{\max}, c_{\max}]$, we have that*

$$\mathbb{E}_\nu \left[\Psi(\mathcal{E}, e) \right] \leq \mathbb{E}_\nu \left[\Psi(\mathcal{E}, e') \right].$$

(ii) Regardless of the distribution of error terms $(e_i)_{i \in I}$,

$$\mathbb{E}_\nu[\Psi(\mathcal{E}, e)] \leq |I| c_{\max} \sqrt{\frac{2\alpha}{|D_0|}} \left(1 + \sqrt{\log M} \right).$$

[13] The fact that $\alpha \geq 1$ follows from the observation that $\alpha = \mathbb{E}_{\mu_I} \left[{}^{\mu_I(\tau)}/_{\mu_{D_0}(\tau)} \right] \geq {}^1/_{\mathbb{E}_{\mu_I} \left[{}^{\mu_{D_0}(\tau)}/_{\mu_I(\tau)} \right]} = 1$.

Sparse linear classifiers. It is informative to evaluate the bounds provided in Proposition 4 for a natural class of partitions $\mathcal{E}$: those generated by sparse linear classifiers. Specifically, we assume that type space T is a subset of $\mathbb{R}^f$ (we will use the inequality $f \leq |T| \leq |D_0|$). For $d \in \{2, \cdots, f\}$, a d-sparse vector $v = (v_k)_{k \in \{1, \cdots, f\}} \in \mathbb{R}^f$ is a vector with at most d non-zero coordinates. The family of partitions $\mathcal{E}$ induced by d-sparse classifiers is defined as

$$\begin{aligned} \mathcal{E} \equiv & \{E_v \equiv \{\eta_v^+, \eta_v^-\} | v \in \mathbb{R}^f, \ v \, d\text{-sparse}\} \\ & \text{where } \eta_v^+ = \{\tau \in T \ s.t. \ \langle \tau, v \rangle > 0\} \text{ and } \eta_v^- = \{\tau \in T \ s.t. \ \langle \tau, v \rangle < 0\}. \end{aligned}$$

The private plan is allowed to use any $d-$sparse linear classifier to decide whether or not to select a particular set of types or not.

Corollary 1. *When possible selection partitions $\mathcal{E}$ are those induced by all d-sparse classifiers, the maximum expected loss $\mathbb{E}_\nu [\Psi(\mathcal{E}, e)]$ from strategic capitation satisfies*

$$\mathbb{E}_\nu [\Psi(\mathcal{E}, e)] \leq 4 c_{\max} |I| \sqrt{\frac{\alpha d \log |D_0|}{|D_0|}}. \tag{18}$$

Indeed, the number of possible partitions of $|T|$ points generated by d-sparse linear classifiers is bounded by $2^d \cdot \binom{f}{d} \cdot \binom{|T|}{d} < \frac{1}{4}|T|^{3d}$, where $\binom{m}{n} = \frac{m!}{(m-n)!n!}$.[14] Since each $E \in \mathcal{E}$ contains two elements, we obtain that $M \leq K^{2d}$. Corollary 1 follows from a direct application of Lemma 1 and the fact that $|T| \leq |D_0|$.

Note that for all practical purposes, term $\sqrt{\log |D_0|}$ may be treated as a constant between 4 and 5. Indeed, for $|D_0| = 48 \times 10^6$, approximately the size of the US Medicare population, $\sqrt{\log |D_0|} \simeq 4.2$, while for $|D_0| = 7 \times 10^9$, roughly the current world population, $\sqrt{\log |D_0|} \simeq$

[14] To obtain this bound, observe that there are $\binom{f}{d}$ ways to choose the d non-zero coordinates in the d-sparse classifier. For each such choice, the classifier can be written in the form $a_1 x_1 + \ldots + a_d x_d < 1$, where $x_1, \ldots, x_d$ are the relevant coordinates, and $a_1, \ldots, a_d \in \mathbb{R}$ are appropriately chosen coefficients. The set of appropriate d-tuples $(a_1, \ldots, a_d)$ forms a polytope $\mathcal{A}$ in $\mathbb{R}^d$, with each of the $|T|$ points representing a linear constraint on the possible values of $(a_1, \ldots, a_d)$. A node of such a polytope is an intersection of d constraints, and thus $\mathcal{A}$ can be identified using d points from T along with the signs of the d constraints. This gives at most $\binom{|T|}{d} \cdot 2^d$ choices.

4.8.

4.2 Unimprovability of Strategic Capitation

In the spirit of Hartline and Roughgarden (2008), we now provide a lower-bound for the minimal efficiency losses that any mechanism can guarantee. Following the notation of Section 2, a state of the world is described by a tuple

$$\omega = (c(\tau, p), K(\cdot), D_0, D_1, H)_{\substack{p\in\{p_0,p_1\}\\ \tau\in T}} \in \Omega,$$

consisting of a distribution of treatment costs $c(\tau, p)$ conditional on types and plan, selection costs K for the private plan, data sets D_0 and D_1 for the public and private plan, as well as hold-out data H privately observed by the public plan.

State of the world ω is drawn according to common prior $\nu \in \Delta(\Omega)$. To provide lower bounds on worst case efficiency losses, it is sufficient for us to consider the class of priors such that sample size $|D_0|$ and distributions of types $\mu_I \in \Delta(T)$ and public data $\mu_{D_0} \in \Delta(T)$ are known.

We consider the problem of Bayes-Nash implementation using budget-balanced direct mechanisms g of the following form:

- data D_0 is publicly observable;
- plan p_1 sends a message $m_1 = (D_1^m, K^m(\cdot)) \in \nu_{|D_1,K(\cdot)}$, reporting her data and selection costs;
- the mechanisms suggests a selection $\lambda_g(D_0, m_1) \in [0,1]^T$ by private plan p_1;
- plan p_1 makes a selection decision $\lambda \in [0,1]^T$, with realized selection $\Lambda \subset I$;
- plan p_0 sends a message $m_0 = H_R \in \operatorname{supp} \nu_{|H}$ corresponding to a reported hold-out sample;
- transfers $\Pi(D_0, m_1, m_0, \Lambda)$ from p_0 to p_1 are implemented.

We denote by $\mathcal{G}_\nu$ the set of incentive compatible direct revelation mechanisms under prior ν. For any direct revelation mechanism $g \in \mathcal{G}_\nu$, the surplus $S(g,\nu)$ attained by mechanism g under prior ν is

$$S(g,\nu) = \mathbb{E}_\nu \left[\sum_{i \in \Lambda} \kappa(p_0, \tau_i) - \kappa(p_1, \tau_i) \,\middle|\, \lambda_g \right] - K(\lambda_g).$$

In turn, given a class $\mathcal{E}$ of partitions, the efficiency loss $L_{\mathcal{E}}(g,\nu)$ of mechanism g relative to treatment allocations measurable with respect to $E \in \mathcal{E}$ is defined as:

$$L_{\mathcal{E}}(g,\nu) = \mathbb{E}_\nu \left[\max_{E \in \mathcal{E}} S_{E|\mathcal{D}_0,\mathcal{D}_1} - S(g,\nu) \right].$$

The following lower bound on efficiency holds.

Proposition 5. *There exists $k > 0$ such that for any class of partitions $\mathcal{E}$,*

$$\max_{\nu} \min_{g \in \mathcal{G}_\nu} L_{\mathcal{E}}(\nu, g) \geq k|I| c_{max} \max_{E \in \mathcal{E}} \mathbb{E}_{\mu_I} \left[\frac{1}{\sqrt{|D_0^\eta|}} \right]. \tag{19}$$

In particular, the efficiency loss achieved by strategic capitation for linear classifiers (Corollary 1) is tight up to an order $\sqrt{\log |D_0|}$, which, for all plausible values of $|D_0|$, can be treated as a constant less than 5.

5 Discussion

This paper explores the value of Big Data in reducing the extent of adverse selection in government-run capitation schemes. We argue that at the correct Big Data limit, including an increasing number of covariates as part of an ex ante capitation formula is unlikely to succeed. Instead we suggest that Big Data may be used to align incentives by using ex post capitation adjustments that interact an unbiased estimate of counterfactual costs to the public plan, with the private plan's deviation from legitimate selection.

This section discusses additional extensions, including the use of strategic capitation in exchanges, as well as dealing with dynamic selection, risk-inflation, and heterogeneity in the quality of care.

5.1 Adverse Selection in Exchanges

Adverse selection is a significant concern in insurance markets such as the ones organized by American Healthcare Act. Indeed, if regulation constrains prices to depend only on a subset of observables (as is the case with community rating), plans will have incentives to select patients that are cheaper to serve given characteristics excluded from legal pricing formulas. This increases the cost of serving patients and can result in limited entry. A simple example suggests that strategic capitation may help improve market outcomes in such environments.

A stylized model. As in Section 2, a set I of patients with types $\tau \in T$ has inelastic unit demand for insurance, where insurance corresponds to a single standardized insurance contract. Plan p_0 is now an incumbent private plan, while p_1 is a potential entrant. For simplicity, we assume that each plan's cost technology is the same: $\forall \tau \in T$, $c(p_0, \tau) \sim c(p_1, \tau)$. Here the objective is not to improve the allocation of patients to plans, but rather to increase competition so that insurance is priced at marginal cost. By law, plans are constrained to offer prices $\pi(\eta)$ that depend only on a coarse set of patient characteristics $\eta \in E$, where E is a partition of T. Prices are bounded above by $\overline{\pi}$.[15]

We assume that the private plans both know their common expected cost of treatment $\kappa(\tau)$ conditional on type τ. Let $\kappa(\eta) \equiv \mathbb{E}_{\mu_I}[\kappa(\tau)|\eta]$. Each plan p has access to a hold-out sample of its own cost H_p. We assume that both plans have lexicographic preferences over maximizing their own revenue and minimizing that of their competitor. The timing of decisions is as follows:

1. potential entrant p_1 decides to enter the market or not;

[15] Parameter $\overline{\pi}$ may be viewed as the patients' (common) value for insurance.

2. each plan p active in the market submits a price formula $\pi_p : \eta \mapsto \pi_p(\eta)$;

3. each plan p active in the market attempts to select a distribution λ_p of patients;

4. if $\pi_{p_0}(\eta) \neq \pi_{p_1}(\eta)$, patients of type η purchase insurance from the cheapest plan; if $\pi_{p_0}(\eta) = \pi_{p_1}(\eta)$, plan p serves distribution of patients $\lambda_p + [\frac{\mu_I}{2} - \lambda_{\neg p}]$, where $\neg p$ denotes the other plan.[16]

The cross-price elasticity of patient demand is infinite, so that patients always go to the cheapest plan. As a result an entrant will at most make zero profit when entering. We assume that whenever the entrant can guarantee itself zero profits it enters.[17] The cost of engaging in selection λ_p is denoted by $K(\lambda_p)$. We assume that K is strictly convex, continuously differentiable, and minimized at $\lambda_p = \frac{\mu_I}{2}$. We denote by Λ_p the realized selected sample of patients purchasing from plan p.

The following result holds.

Proposition 6. *The market entry game described above has a unique subgame perfect equilibrium in which the potential entrant does not enter, and the incumbent charges price* $\pi_{p_0}(\eta) = \overline{\pi}$.

In the off-equilibrium subgame following entry both the entrant and the incumbent make equilibrium losses $-K(\lambda^*) < 0$ *where* λ^* *solves* $\max_{\lambda \in [0,1]^T} \left[\sum_{\tau \in T} \lambda(\tau) \left(\kappa(\eta) - \kappa(\tau)\right) - K(\lambda)\right]$.

Indeed, because cross-price elasticities are infinite, in equilibrium, both plans price at marginal cost conditional on η: $\pi_p(\eta) = \kappa(\eta)$. Furthermore, since the marginal cost of selection at $\lambda_p = \mu_I/2$ is zero, both players find it profitable to engage in non-zero selection. In aggregate however, selection efforts cancel one another and merely destroy surplus.

Strategic capitation. Consider now the following extension of the strategic capitation scheme introduced in Section 3. The game described above is modified in two ways:

[16] We assume that the cost of selection $K(\lambda_p)$ is sufficiently steep around $\frac{\mu_I}{2}$ that $\lambda_p + \frac{\mu_I}{2} - \lambda_{\neg p} \in \Delta(T)$ for all individually rational selection policies.

[17] This could be due to small subsidies for entry, or high but finite cross-price elasticities.

- at stage 2, along with submitting pricing formulas $\pi_p(\cdot)$, each active plan submits a report $H_{R,p}$ of their hold-out sample.

- after selection has occurred, for each type η it serves, plan p receives price $\pi_p(\eta)$ and capitation adjustment $\Delta\pi(\eta, H_{R,\neg p}, \Lambda_p)$ taking the form:

$$\Delta\pi(\eta_i, H_{R,p}, \Lambda_p) \equiv \text{cov}_I(s_{i,p}, r_{i,p}|\eta_i = \eta) = \frac{1}{|I^\eta|}\sum_{i \in I^\eta} s_{i,p} r_{i,p},$$

  with

$$s_{i,p} \equiv \frac{\mu_{\Lambda_p}(\tau_i|\eta_i)}{\mu_I(\tau_i|\eta_i)} - 1 \quad \text{and} \quad r_{i,p} \equiv \frac{1}{|H^{\tau_i}_{R,\neg p}|}\sum_{j \in H^{\tau_i}_{R,\neg p}} \left[\widehat{c}^{R,\neg p}_j(p_0) - \pi_p(\eta)\right].$$

Proposition 7. *The market game with strategic capitation described above has an efficient truthful equilibrium in which: the potential entrant enters; both plans submit prices* $\pi_p(\eta) = \kappa(\eta)$*; both plans select a representative population in expectation (*$\lambda_p = \mu_I/2$*); both plans submit their hold-out sample costs truthfully (*$H_{R,p} = H_p$*); expected ex post adjustments are equal to 0 (*$\mathbb{E}\Delta\pi_p = 0$*).*

The intuition for this result is identical to that of Proposition 3. Given ex ante representative selection, a plan's expected capitation adjustment is equal to zero regardless of messages sent by the other plan. Given truthful revelation of costs, representative selection is a best-response.

5.2 Extensions and Implementation Concerns

Dynamic Selection and Risk-Inflation. The process of selection is dynamic. In the context of Medicare Advantage, patients have the opportunity to switch back and forth between public and private plans once a year. This implies that costs of care need to be evaluated over time. Plans with low short-term cost of care may end up generating greater longer term costs if they skimp on quality, and encourage patients to disenroll once they get sick enough (Ellis, 1998). Appendix A shows how to adjust strategic capitation to address this issue. It becomes important to keep track of the counterfactual distribution of types, should the patient have remained with the public plan.

One noteable insight from Appendix A is that correct dynamic capitation fees remove concerns over risk-inflation by private plans. Indeed, if a patient with legitimate characteristic η_t enrolls in the private plan at time t, then baseline repayments π_{t+s} to the private plan at all times $t+s$ where the patient remains with the private plan take the form

$$\pi_{t+s} = \pi(t+s, \eta_t) \equiv \mathbb{E}[\widehat{c}_{i,t+1}(p_0)|\eta_t].$$

In other words, target repayments depend only on the type η_t of the patient when she enrolls with the private plan, and on elapsed time $t+s$. *It does not depend on the patient's type* η_{t+s} *after enrollment time* t. As a result, the plan has no incentives to exaggerate the medical condition of patients it enrolls (for instance by running a battery of tests detecting mild conditions). This is not the case when target repayments π_{t+s} depend on types η_{t+s} at time $t+s$.

Quality. Throughout the paper we assume that the quality of actual healthcare delivery is homogeneous across plans. In practice, insurance plans may differ in the quality of care they deliver to their enrollees. It is important to take into account such quality outcomes when designing capitation schemes. If not, costs may be kept low at the expense of quality. Appendix A describes an extension of strategic capitation that correctly reflects differences in the quality of care. An important limitation is that it requires that health outcomes (including death) be observable, and that they be assigned monetary values.

Surplus Extraction. The paper focuses on the efficient allocation of patients across public and private plans. However, if there is a deadweight loss to public funds, it may be welfare improving for the public plan to extract some of the surplus. Since the private plans' has private information over her costs conditional on patient types, this is a difficult multidimensional screening problem. Two observations are helpful to make progress on this issue. First, given that we consider prior-free mechanisms, the argument of Carroll (2015) suggests there may not be much value in complex multidimensional screening. It may be near-optimal to focus on separable one-dimensional screening mechanisms that associate a discounted baseline capitation rate $\rho(\eta)\kappa(\eta, p_0)$ with $\rho(\eta) \in [0, 1]$ to each patient with characteristics η. A second useful observation is that strategic capitation adjustments used to prevent selection of mispriced types can be applied to any baseline repayment scheme. This suggests using capitation schemes of the form

$$\Pi(\Lambda, \tau_I, H_R) \equiv \sum_{i \in \Lambda} \rho(\eta)\kappa(\eta, p_0) + \Delta\pi(\eta_i, H_R)$$

where $\rho(\eta) \in [0, 1]$ is a given discounting profile, and $\Delta\pi(\eta_i, H_R) = \frac{1}{|I^\eta|} \sum_{i \in I^\eta} s_i r_i$, with $s_i \equiv \frac{\mu_\Lambda(\tau_i|\eta_i)}{\mu_I(\tau_i|\eta_i)} - 1$ and $r_i \equiv \frac{1}{|H_R^{\tau_i}|} \sum_{j \in H_R^{\tau_i}} \left[\widehat{c}_j^R(p_0) - \kappa(\eta, p_0)\right]$. This separates the problem of extracting revenue, and that of preventing illegitimate selection.

Ethics. Regulators frequently ban indexing ex ante capitation rates on certain observables, such as ethnicity or income. One rationale for this is that the law has expressive content that affects social norms, and it is desirable to reinforce the norm that all citizens deserve equal treatment. This has subtle consequences on the social acceptability of contingent incentive schemes: having different equilibrium capitation payments for citizens of different ethnic background seems repugnant in a sense related to Roth (2007); but punishing discrimination against specific ethnic groups (which should mostly remain off of the equilibrium path) does not. The adjustments proposed by strategic capitation fall in this latter category: it punishes plans for non-representative selection of types.

Volatility of revenues and profits. One concern with strategic capitation is that the capitation payments that the private plan ultimately receives are uncertain at the interim stage: if noise in selection causes the private plan's to enroll types that are relatively cheap (resp. expensive) to treat for the public plan, it receives lower (resp. higher) payments than anticipated. While this increases the volatility of revenues, this may in fact reduce the volatility of profits. Indeed, types that are relatively cheap (resp. expensive) to treat to the public plan are also likely to be cheap (resp. expensive) to treat for the private plan. If noise in selection causes a plan to overselect types that are relatively cheap to treat for the public plan, it receives a negative capitation adjustment. However, it is likely that the cost of treating these types was also relatively cheap for the private plan, keeping net profits stable. A similar reasoning applies if the plan overselects types that are relatively expensive to treat for the public plan. The positive capitation adjustments may well compensate a corresponding increase in the private plan's cost of care. In other words, strategic capitation may serve as insurance against selection shocks.

A Extensions

A.1 Dynamic Selection

In dynamic settings, capitation schemes need to control for differential transitions in health status across plans. For simplicity, as in Section 3, we assume that expected costs conditional on legitimate characteristics are known, and that at each time t, comparative advantage depends on a commonly known set of legitimate selection characteristics $\eta \in E$. We denote by $\tau_{i,t}$ the type of patient i at date t, by $\eta_{i,t}$ her legitimate selection characteristic at date t, and by $\widehat{c}_i(t,p)$ her realized cost of care if treated by plan p at time t. Types $(\tau_t)_{t\in\{0,\cdots,\overline{T}\}}$ and characteristics $(\eta_t)_{t\in\{0,\cdots,\overline{T}\}}$ follow separate Markov chains, summarized under notation Φ_p, which depend on the plan p that the patient is enrolled with. Future costs are discounted using discount factor $\delta \in (0,1]$, and $\overline{T}$ denotes an upper bound to the duration of patients' lives in the system.

For a patient i of type τ_i enrolled with the public plan from time t to time $\overline{T}$, we define

$$\widehat{C}_i(t, p_0) \equiv \sum_{s=t}^{\overline{T}} \delta^{s-t} \widehat{c}_i(s, p_0) \quad \text{and} \quad C(t, \eta, p_0) \equiv \mathbb{E}_\nu \left[\widehat{C}_i(t, p_0) \middle| \eta_t = \eta \right].$$

In dynamic environments, strategic capitation must accommodate the possible reenrollment of patients with the public plan. As a result, transfers must occur at the reentry of patients into the public system. Let us denote by Λ_t the selection of patients enrolled with the private plan at time t, and by Λ_t^{re} the selection of patients disenrolling from the private plan and enrolling with the public plan at time t. The following scheme generalizes strategic capitation. At initial time of enrollment $t = 0$, the public plan commits to the following baseline payments conditional on legitimate characteristics $\eta \in E$:

- a capitation payment $\pi(t, \eta_0) = \mathbb{E}_\nu[\widehat{c}_i(t, p_0) | \eta_{i,0} = \eta_0]$ whenever patient i with initial type η_0 is enrolled with the private plan at time t;
- a signed transfer π_i^{re} (with positive transfers being made from the public plan to the private plan) at every time T such that patient i returns to the public plan: $\pi_i^{re} = \mathbb{E}_\nu[\widehat{C}(T, p_0) | \eta_{i,0}, \Phi_{p_0}] - C(T, \eta_{i,T}, p_0)$.

Provided that the private plan does not engage in illegitimate selection, this scheme induces efficient dynamic behavior by the private plan. To dissuade illegitimate selection, dynamic strategic capitation makes adjustments $\Delta\pi(t, \eta_0)$ and $\Delta\pi^{re}(T, \eta_0)$ using reported hold-out data H_R as follows:

- $\Delta\pi(t, \eta_0) = \frac{1}{|I^{\eta_0}|} \sum_{i \in I^{\eta_0}} s_{i,t} r_{i,t}$, with

$$s_{i,t} \equiv \frac{\mu_{\Lambda_t}(\tau_{i,0} | \eta_{i,0} = \eta_0)}{\mu_I(\tau_{i,0} | \eta_{i,0} = \eta_0)} - 1, \quad \text{and} \quad r_{i,t} \equiv \frac{1}{|H_R^{\tau_{i,0}}|} \sum_{j \in H_R^{\tau_{i,0}}} \widehat{c}_j^R(t, p_0) - \pi(t, \eta_0).$$

- $\Delta\pi^{re}(t,\eta_t) = \frac{1}{|I^{\eta_t}|}\sum_{i\in I^{\eta_t}} s^{re}_{i,t}\, r^{re}_{i,t}$, with

$$s^{re}_{i,t} \equiv \frac{\mu_{\Lambda^{re}_t}(\tau_{i,t}|\eta_{i,t}=\eta_t)}{\mu_I(\tau_{i,t}|\eta_{i,t}=\eta_t)} - 1, \quad \text{and} \quad r^{re}_{i,t} \equiv \frac{1}{|H_R^{\tau_{i,t}}|}\sum_{j\in H_R^{\tau_{i,t}}}\left[C(T,\eta_{i,T},p_0) - \widehat{C}^R_j(t,p_0)\right].$$

A.2 Quality

If the private and public plan differ in the quality of health outcomes they deliver to patients, the value associated with different health outcomes needs to be reflected in capitation transfers. We assume that health outcomes (including death) for each patient $i \in I$ treated by plan p are observable and associated with realized monetary values $\widehat{v}_i(p)$. By analogy to costs, we assume that the private plan's advantage function is measurable with respect to a relatively small set of types η_i. The strategic capitation scheme can then be extended to the scenario with outcome qualities. Given selection rule λ and transfers Π, the surpluses accruing to the public and private plans take the form

$$\begin{aligned}
\mathbb{E}_\nu U_0 &= \mathbb{E}_\nu\left[-\Pi + \sum_{i\in\Lambda}\widehat{c}_i(p_0) + \widehat{v}_i(p_1) - \widehat{v}_i(p_0)\middle|\lambda\right],\\
\mathbb{E}_\nu U_1 &= \mathbb{E}_\nu\left[\Pi - \sum_{i\in\Lambda}\widehat{c}_i(p_1)\middle|\lambda\right] - K(\lambda).
\end{aligned}$$

Differences in quality of care are isomorphic to a change in the public plan's cost of care. Since we assume that health outcomes are observable, data D_0 should now include values $\widehat{v}_i(p_0)$ to patients in D_0, and health outcomes $\widehat{v}_i(p_1)$ to patients in D_1 should be visible to the public plan. Strategic capitation can be extended by setting transfers:

$$\Pi(\Lambda,\tau_I,H_R) \equiv \sum_{i\in\Lambda}\widehat{v}_i(p_1) + \pi(\eta_i) + \Delta\pi(\eta_i,H_R)$$

where

$$\pi(\eta) \equiv \sum_{\tau\in\eta} \mu_I(\tau|\eta) \left[\frac{1}{|D_0^\tau|} \sum_{i\in D_0^\tau} \widehat{c}_i(p_0) - \widehat{v}_i(p_0) \right]$$

and $\Delta\pi(\eta, H_R)$ takes the form:

$$\Delta\pi(\eta_i, H_R) \equiv \frac{1}{|I^\eta|} \sum_{i\in I^\eta} s_i r_i,$$

with $s_i \equiv \frac{\mu_\Lambda(\tau_i|\eta_i)}{\mu_I(\tau_i|\eta_i)} - 1$ and

$$r_i \equiv \frac{1}{|H_R^{\tau_i}|} \left[\sum_{j\in H_R^{\tau_i}} \widehat{c}_j^R(p_0) - \widehat{v}_j(p_0) \right] - \pi(\eta).$$

B Proofs

B.1 Proofs for Section 3

Proof of Proposition 1: We begin with point (i). Reports from plan p_0 do not affect reimbursements so that truth-telling strategy β^* is dominant. In turn, for any selection Λ measurable with respect to characteristics $\eta \in E$, the private plan's expected payoffs from selection take the form

$$\mathbb{E}_c \left[\sum_{i\in I} \mathbf{1}_{i\in\Lambda}(\kappa(\eta_i, p_0) - \kappa(\eta_i, p_1) + e_{i,\tau_i}) \right] = \mathbb{E}_c \left[\sum_{i\in I} \mathbf{1}_{i\in\Lambda}(\kappa(\eta_i, p_0) - \kappa(\eta_i, p_1)) \right]$$

where we used the fact that $\mathbb{E}_c[e_{i,\tau_i}|\eta_i] = 0$. It follows that the optimal selection rule is indeed $\Lambda = \Lambda_{\max} \equiv \{i \mid \kappa(\eta_i, p_0) - \kappa(\eta_i, p_1) > 0\}$.

Let us turn to point (ii). It is useful to define

$$\xi_\eta \equiv \log\left(1 + \frac{\kappa(\eta, p_0) - \kappa(\eta, p_1)}{\underline{\kappa}}\right) + \frac{1}{2(|D_0^\tau| + 2)},$$

where we use the convention that $\log(x) = -\infty$ for $x \leq 0$.

Given data D_0 and signal x_τ the private plan's conditional belief over random cost parameter ε_τ follows a normal distribution $\mathcal{N}(\chi_\tau, \sigma_\tau^2)$ with

$$\chi_\tau = \frac{x_\tau + \sum_{i \in D_0^\tau} x_i}{2 + |D_0^\tau|} \quad \text{and} \quad \sigma_\tau^2 = \frac{1}{2 + |D_0^\tau|}.$$

This implies that $\mathbb{E}_\nu[e_{i,\tau}|D_0^\tau, x_\tau] = \underline{\kappa}\left[\exp\left(\chi_\tau - \frac{1}{2(|D_0^\tau|+2)}\right) - 1\right]$. Furthermore, conditional on getting a data set of cardinal $|D_0^\tau|$, posterior belief χ_τ itself follows a Gaussian distribution $\mathcal{N}\left(0, \frac{(|D_0^\tau|+1)^2 + |D_0^\tau|}{(|D_0^\tau|+2)^2}\right)$.

We prove the first part of (ii) by showing that

$$\begin{aligned} S_{\max} - S^{\mathsf{sparse}} = & \sum_{i \in I} \operatorname{prob}\left(\chi_{\tau_i} \geq \xi_{\eta_i}\right)\left[\kappa(\eta_i, p_0) - \kappa(\eta_i, p_1)\right]^+ \\ & + \sum_{i \in I} \operatorname{prob}\left(\chi_{\tau_i} \leq \xi_{\eta_i}\right)\left[\kappa(\eta_i, p_1) - \kappa(\eta_i, p_0)\right]^+. \end{aligned}$$

We prove the second part of (ii) by showing that if the private plan has no comparative advantage, i.e. $\kappa(\eta, p_0) = \kappa(\eta, p_1)$, the public plan makes losses

$$\mathbb{E}_\nu[U_0|D_0, x, \Pi^{\mathsf{sparse}}] = -\underline{\kappa} \sum_{i \in I}\left[\exp\left(\chi_{\tau_i} - \frac{1}{2(|D_0^{\tau_i}| + 2)}\right) - 1\right]^- .$$

Indeed, conditional on her information (x_τ, D_0^τ), plan p_1's expected payoff from selecting a patient of type τ is

$$\kappa(\eta, p_0) - \kappa(\eta, p_1) - \mathbb{E}_\nu[e_{i,\tau}|D_0^\tau, x_\tau].$$

Since $\mathbb{E}_\nu[e_{i,\tau}|D_0^\tau, x_\tau] = \underline{\kappa}\left[\exp\left(\chi_\tau - \frac{1}{2(|D_0^\tau|+2)}\right) - 1\right]$, plan p_1 will select type τ if and only if

$$\kappa(\eta, p_0) - \kappa(\eta, p_1) - \underline{\kappa}\left[\exp\left(\chi_\tau - \frac{1}{2(|D_0^\tau| + 2)}\right) - 1\right] > 0 \iff \chi_\tau < \xi_\eta.$$

This implies that efficiency losses indeed take the form

$$L^{\mathsf{sparse}} = \sum_{i \in I} \text{prob}\left(\chi_{\tau_i} \geq \xi_{\eta_i}\right) \left[\kappa(\eta_i, p_0) - \kappa(\eta_i, p_1)\right]^+ + \sum_{i \in I} \text{prob}\left(\chi_{\tau_i} \leq \xi_{\eta_i}\right) \left[\kappa(\eta_i, p_1) - \kappa(\eta_i, p_0)\right]^+.$$

When plan p_1 has no comparative advantage, it selects all types τ such that $\mathbb{E}_\nu[e_{i,\tau}|D_0^\tau, x_\tau] < 0$, and p_0's expected payoffs are equal to

$$\begin{aligned}\mathbb{E}_\nu[U_0|D_0, D_1] &= \sum_{i \in I} \mathbb{E}_\nu[e_{i,\tau_i}|D_0, D_1] \mathbf{1}_{\mathbb{E}_\nu[e_{i,\tau}|D_0,x]<0} \\ &= -\underline{\kappa} \sum_{i \in I} \left[\exp\left(\chi_{\tau_i} - \frac{1}{2(|D_0^{\tau_i}|+2)}\right) - 1\right]^- .\end{aligned}$$

□

Proof of Proposition 2: It is useful to define

$$\forall \tau \in T, \quad \zeta_\tau \equiv \log\left(1 + \frac{\kappa(\eta, p_0) - \kappa(\eta, p_1)}{\underline{\kappa}} \exp\left(-\chi_{p_0,\tau} + \frac{1}{2(|D_0^\tau|+2)}\right)\right).$$

Plan p_1's expected profit from selecting a patient of type τ is

$$\begin{aligned}&\kappa(\eta, p_0) - \kappa(\eta, p_1) + \mathbb{E}_\nu[e_{i,\tau}|D_0^\tau] - \mathbb{E}_\nu[e_{i,\tau}|D_0^\tau, x_\tau] \\ &\quad = \kappa(\eta, p_0) - \kappa(\eta, p_1) + \underline{\kappa}\left[\exp\left(\chi_{p_0,\tau} - \frac{1}{2(|D_0^\tau|+1)}\right) - \exp\left(\chi_{p_1,\tau} - \frac{1}{2(|D_0^\tau|+2)}\right)\right].\end{aligned}$$

This implies that plan p_1 will select patients of type τ if and only if[18]

$$\begin{aligned} &\zeta_\tau > \chi_{p_1,\tau} - \chi_{p_0,\tau} + \frac{1}{2(|D_0^\tau|+1)(|D_0^\tau|+2)} \\ \iff &\zeta_\tau > \frac{1}{|D_0^\tau|+2}\left(x_\tau - \chi_{p_0,\tau} + \frac{1}{2(|D_0^\tau|+1)}\right). \end{aligned}$$

Observing that ζ_τ has the same sign as $\kappa(\eta, p_0) - \kappa(\eta, p_1)$, this implies that the efficiency loss L^{rich} can be written as

$$\begin{aligned} L^{\mathsf{rich}} = &\sum_{i\in I} \mathrm{prob}_{x_{\tau_i}}\left(x_{\tau_i} - \chi_{p_0,\tau} + \frac{1}{2(|D_0^\tau|+1)} < -(|D_0^\tau|+2)\zeta_{\tau_i}^-\right)[\kappa(\eta_i,p_0)-\kappa(\eta_i,p_1)]^- \\ &+ \sum_{i\in I} \mathrm{prob}_{x_{\tau_i}}\left(x_{\tau_i} - \chi_{p_0,\tau} + \frac{1}{2(|D_0^\tau|+1)} > (|D_0^\tau|+2)\zeta_{\tau_i}^+\right)[\kappa(\eta_i,p_0)-\kappa(\eta_i,p_1)]^+ \end{aligned}$$

where we use the convention $z^- = \max\{0, -z\}$. The first term corresponds to the inefficiency loss from types that are more efficiently treated by p_0 but end up selected by p_1. The second term corresponds to the inefficiency loss from types that are more efficiently treated by p_1, but end up being treated by p_0.

Recall that $\chi_{p_0,\tau} \sim \mathcal{N}\left(0, \frac{|D_0^\tau|^2+|D_0^\tau|}{(1+|D_0^\tau|)^2}\right)$, and therefore there are constants $c_1, c_2, c_3 > 0$ such that with probability greater than $1/2$,

$$c_1 \cdot \frac{|\kappa(\eta_i,p_0)-\kappa(\eta_i,p_1)|}{\underline{\kappa}} \le |\zeta_{\tau_i}| \le c_2 \cdot \frac{|\kappa(\eta_i,p_0)-\kappa(\eta_i,p_1)|}{\underline{\kappa}}, \tag{20}$$

and for all $t > 0$, the probability that $|\zeta_{\tau_i}| < \exp(-t) \cdot |\kappa(\eta_i, p_0) - \kappa(\eta_i, p_1)|/\underline{\kappa}$ is at most $\exp(-c_3 t^2)$.

For the upper bound (9), suppose that $|\kappa(\eta_i, p_0) - \kappa(\eta_i, p_1)|/\underline{\kappa} = s/|D_0^\tau|$ for some $s > 1$.

[18] Selection will not occur when ζ_τ is not defined.

We have,

$$
\begin{aligned}
\operatorname{prob}_{x_{\tau_i},\chi_{p_0,\tau}}\left(x_{\tau_i}-\chi_{p_0,\tau}+\frac{1}{2(|D_0^\tau|+1)}>(|D_0^\tau|+2)\zeta_{\tau_i}^+\right)&<\\
\operatorname{prob}_{x_{\tau_i},\chi_{p_0,\tau}}\left(x_{\tau_i}-\chi_{p_0,\tau}>\sqrt{s}\right)+\operatorname{prob}_{\chi_{p_0,\tau}}\left(\zeta_{\tau_i}^+<\frac{\sqrt{s}}{|D_0^\tau|}\right)&<\\
\exp(-c_4\cdot s^2)+\operatorname{prob}_{\chi_{p_0,\tau}}\left(|\zeta_{\tau_i}|<s^{-1/2}\cdot|\kappa(\eta_i,p_0)-\kappa(\eta_i,p_1)|/\underline{\kappa}\right)&\leq\\
\exp(-c_4\cdot s^2)+\exp(-c_3(\log s)^2)&<\\
\frac{c_5}{s}&,
\end{aligned}
$$

for some constants $c_4, c_5 > 0$. Therefore, the expected contribution of patient $i \in I$ to efficiency loss $S_{\max} - S^{\mathsf{rich}}$ is bounded above by

$$
\frac{c_5}{s}\cdot\frac{|\kappa(\eta_i,p_0)-\kappa(\eta_i,p_1)|}{\underline{\kappa}}\cdot\underline{\kappa}=\frac{c_5}{|D_0^\tau|}\cdot\underline{\kappa}.
$$

When $s \leq 1$, the contributions of i to efficiency loss is bounded above by $|\kappa(\eta_i, p_0) - \kappa(\eta_i, p_1)| \leq \underline{\kappa}/|D_0^\tau|$, thus completing the proof of the upper bound.

We now prove the lower bound (10). For concision, we use the notation $\delta \equiv \mathbb{E}_{\mu_I}\frac{1}{|D_0^\tau|} \in (0, 1]$. We will prove an efficiency loss of $c_6\delta^2\underline{\kappa}$ for some $c_6 > 0$. We first claim that there exists $k > 0$ such that $\operatorname{prob}_{\mu_I}(|D_0^\tau| \leq k) \geq \delta^2 k/10$. Suppose this is not the case. Then

$$
\begin{aligned}
\delta=\mathbb{E}_{\mu_I}\frac{1}{|D_0^\tau|}&=\sum_{k=1}^{\infty}\frac{1}{k\cdot(k+1)}\cdot\operatorname{prob}_{\mu_I}(|D_0^\tau|\leq k)\\
&\leq\sum_{k=1}^{10/\delta^2}\frac{1}{k\cdot(k+1)}\cdot\operatorname{prob}_{\mu_I}(|D_0^\tau|\leq k)+\sum_{k\geq 10/\delta^2}\frac{1}{k\cdot(k+1)}\\
&<\frac{\delta^2}{10}+\sum_{k=1}^{10/\delta^2}\frac{\delta^2}{10(k+1)}<\frac{\delta^2}{10}+\int_1^{10/\delta^2}\frac{\delta^2}{10x}\mathrm{d}x.
\end{aligned}
$$

Using the fact that $\int_1^{10/\delta^2}\frac{\delta^2}{10x}\mathrm{d}x = \frac{\delta^2}{10}\log\left(\frac{10}{\delta^2}\right) < \frac{\delta^2}{5}\log(4/\delta) < \frac{\delta^2}{5}(4/\delta - 1)$, we obtain a contradiction.

Set $|\kappa(\eta_i, p_0) - \kappa(\eta_i, p_1)| = \underline{\kappa}/k$. By our choice of k, for a fraction of patients greater than $\delta^2 k/10$, $|D_0^\tau| \leq k$. By (20), with probability greater than 1/2, $|\zeta_{\tau_i}| \leq c_2/k$. Thus $|\zeta_{\tau_i}| \cdot (|D_0^\tau|+2) < c_2+2$, and for some $\epsilon > 0$, i contributes at least ϵ efficiency loss $S_{\max} - S^{\mathsf{rich}}$ with probability greater than ϵ. This leads to a per-patient expected efficiency loss of order

$$c_7 \cdot \frac{\delta^2 k}{10} \cdot \frac{\underline{\kappa}}{k} = (c_7/10) \cdot \delta^2 \underline{\kappa}.$$

When private plan p_1 has no comparative advantage, expected payoffs to the public plan take the form

$$\begin{aligned}
\mathbb{E}_\nu[U_0|D_0, x, \Pi^{\mathsf{rich}}] &= -\underline{\kappa} \sum_{i \in I} \left[\exp\left(\chi_{p_0,\tau_i} - \frac{1}{2(|D_0^{\tau_i}|+1)} \right) - \exp\left(\chi_{p_1,\tau_i} - \frac{1}{2(|D_0^{\tau_i}|+2)} \right) \right]^+ \\
&= -\underline{\kappa} \sum_{i \in I} \exp\left(\chi_{p_0,\tau} - \frac{1}{2(|D_0^{\tau_i}|+1)} \right) \left[1 - \exp\left(\frac{1}{|D_0^{\tau_i}|+2} \left[x_{\tau_i} - \chi_{p_0,\tau_i} + \frac{1}{2(|D_0^{\tau_i}|+1)} \right] \right) \right]^+,
\end{aligned}$$

which is at least of the order of $\underline{\kappa}\, |I| \mathbb{E}_{\mu_I} \frac{1}{|D_0^{\tau_i}|}$ with probability bounded away from zero. $\square$

Proof of Proposition 3: Plan p_1's payoff takes the form

$$\mathbb{E}_\nu \left[\Pi(\Lambda) - \sum_{i \in \Lambda} \widehat{c}_i(p_1) \right] = \mathbb{E}_\nu \left[\sum_{i \in \Lambda} \kappa(\eta_i, p_0) - \kappa(\eta_i, p_1) \right] + \mathbb{E}_\nu \left[\sum_{i \in \Lambda} \Delta\pi(\eta_i, H_R, \Lambda) - e_{i,\tau_i} \right].$$

In any equilibrium (λ, β), the expected cost of transfers to the public plan must be weakly lower under β than under truthful reporting β^*. Recalling that $r_\tau \equiv \frac{1}{|H_R^\tau|} \sum_{j \in H_R^\tau} \left[\widehat{c}_j^R(p_0) - \widehat{\kappa}(\eta, p_0) \right]$ denotes reported residuals from the baseline capitation formula on hold-out sample costs, this implies that

$$\begin{aligned}
\mathbb{E}_\nu \left[\sum_{i \in \Lambda} \Delta\pi(\eta_i, H_R, \Lambda) - e_{i,\tau_i} \middle| \lambda, \beta \right] &= \mathbb{E}_\nu \left[\sum_{\eta \in E} |\Lambda^\eta| \left(\sum_{\tau \in T_\eta} [\mu_\Lambda(\tau|\eta) - \mu_I(\tau|\eta)] \widehat{r}_\tau - \sum_{i \in \Lambda} e_{i,\tau_i} \right) \middle| \lambda, \beta \right] \\
&\leq -\mathbb{E}_\nu \left[\sum_{\eta \in E} |\Lambda^\eta| \sum_{\tau \in \eta} \mu_I(\tau|\eta)] \mathbb{E}_\nu[e_{i,\tau}|\tau] \right] = 0.
\end{aligned}$$

Therefore it follows that plan p_1 gets a payoff at most equal to surplus

$$\mathbb{E}_\nu \left[\sum_{i \in \Lambda} \kappa(p_0, \eta_i) - \kappa(p_1, \eta_i) \right].$$

Since strategic-capitation adjustments have mean to zero when the private plan uses legitimate selection, the private plan can guarantee herself this payoff by using efficient selection strategy $\Lambda_{\max}$. Hence, in any equilibrium $\lambda = \Lambda_{\max}$. Since the private plan uses a legitimate selection rule, the public plan cannot reduce capitation payments by biasing reports, and uses truthful reporting strategy β^*. $\square$

B.2 Proofs for Section 4

Proof of Proposition 4: Let $\kappa(\eta, p) \equiv \mathbb{E}_{\mu_I}[\kappa(\tau, p)|\eta]$ denote the expected cost of service for plan p conditional on legitimate selection characteristic η. Given a partition E and a selection rule λ, plan p_1's expected returns are

$$\begin{aligned}
\mathbb{E}_\nu[U_1|D_0, D_1] =& \mathbb{E}_\nu \left[\sum_{i \in I} \lambda(\tau_i) \left[\widehat{\kappa}(\eta_i, p_0) + \Delta\pi(\eta_i, H_R) - \kappa(\tau_i, p_1) \right] \Big| D_0, D_1 \right] - K(\lambda) \\
=& \mathbb{E}_\nu \left[\sum_{i \in I} \lambda(\tau_i) \left[\kappa(\tau_i, p_0) - \kappa(\tau_i, p_1) \right] \Big| D_0, D_1 \right] - K(\lambda) \\
+& \mathbb{E}_\nu \left[\sum_{i \in I} \lambda(\tau_i) \left[\kappa(\eta_i, p_0) + \Delta\pi(\eta_i, H_R) - \kappa(\tau_i, p_0) \right] \Big| D_0, D_1 \right] \\
+& \mathbb{E}_\nu \left[\sum_{i \in I} \lambda(\tau_i) \left[\widehat{\kappa}(\eta_i, p_0) - \kappa(\eta_i, p_0) \right] \Big| D_0, D_1 \right] \\
=& U_1^A + U_1^B + U_1^C.
\end{aligned}$$

where U_1^A, U_1^B and U_1^C are defined as the three respective terms in the expression above. Note that $U_1^A = S(\lambda|D_0, D_1)$. The key steps of the proof are the following,

(i) in any equilibrium (E, λ, β), $U_1^B \leq 0$;

(ii) for any reporting strategy β, if λ is measurable with respect to E, then $U_1^B = 0$;

(iii) for any E and λ,

$$\left|\mathbb{E}_\nu[U_1^C|D_0]\right| \le \mathbb{E}_\nu[\Psi(\mathcal{E}, e)].$$

Let us first show that points (i), (ii) and (iii) imply properties (15), (16) and (17). We have that under equilibrium strategies (E, λ, β),

$$\begin{aligned}\mathbb{E}_\nu[U_1|D_0, D_1] \le & S(\lambda|D_0, D_1) + \mathbb{E}_\nu[U_1^B|D_0, D_1, \lambda, \beta] + \mathbb{E}_\nu[U_1^C|D_0, D_1, \lambda, \beta] \\ \le & S(\lambda|D_0, D_1) + \mathbb{E}_\nu[U_1^C|D_0, D_1, \lambda, \beta].\end{aligned}$$

In addition, from the fact that the private plan is weakly better off using (E, λ) over any strategy (E', λ') where λ' is measurable with respect to E', it follows that

$$\mathbb{E}_\nu[U_1] \ge \mathbb{E}_\nu\left[\max_{E' \in \mathcal{E}} S_{E'|D_0, D_1}\right] - \mathbb{E}_\nu[\Psi(\mathcal{E}, e)].$$

Overall this implies that $S(\lambda) \ge \mathbb{E}_\nu\left[\max_{E \in \mathcal{E}} S_{E|D_0,D_1}\right] - 2\mathbb{E}_\nu[\Psi(\mathcal{E}, e)]$. Condition (16) follows from the fact that truthful reporting $\beta^*(c, \tau)$ guarantees that

$$\begin{aligned}\mathbb{E}_\nu\left[\sum_{i \in I} \lambda(\tau_i)\left[\widehat{\kappa}(\eta_i, p_0) + \Delta\pi(\eta_i, H_R) - \kappa(\tau_i, p_0)\right]\right] \ge & \underbrace{\mathbb{E}_\nu\left[\sum_{i \in I} \lambda(\tau_i)\left[\kappa(\eta_i, p_0) + \Delta\pi(\eta_i, H_R) - \kappa(\tau_i, p_0)\right]\right]}_{=0} \\ & + \underbrace{\mathbb{E}_\nu\left[\sum_{i \in I} \lambda(\tau_i)\left[\widehat{\kappa}(\eta_i, p_0) - \kappa(\eta_i, p_0)\right]\right]}_{\ge -\mathbb{E}_\nu[\Psi(\mathcal{E}, e)]}.\end{aligned}$$

Finally, condition (17) follows from the fact that plan p_1 can choose a selection strategy measurable with respect to E, which guaranteed p_1 positive expected payoffs.

Let us return to the proofs of points (i), (ii) and (iii) above. Point (i) follows from the fact that in equilibrium the expected transfers of p_0 to plan p_1 under equilibrium reporting strategy β must be weakly lower than under truthful reporting strategy β^*, i.e. $\mathbb{E}_\nu\left[\Pi|\beta\right] \le$

$\mathbb{E}_\nu\left[\Pi|\beta^*\right]$. This implies that

$$\mathbb{E}_\nu\left[\sum_{i\in I}\lambda(\tau_i)\left[\widehat{\kappa}(\eta_i,p_0)+\Delta\pi(\eta_i,H_R)\right]\Big|\beta\right]\leq\mathbb{E}_\nu\left[\sum_{i\in I}\lambda(\tau_i)\left[\widehat{\kappa}(\eta_i,p_0)+\Delta\pi(\eta_i,H_R)\right]\Big|\beta^*\right],\text{ so that}$$

$$\mathbb{E}_\nu\left[\sum_{i\in I}\lambda(\tau_i)\left[\kappa(\eta_i,p_0)+\Delta\pi(\eta_i,H_R)-\kappa(\tau_i,p_0)\right]\Big|\beta\right]\leq\mathbb{E}_\nu\left[\sum_{i\in I}\lambda(\tau_i)\left[\kappa(\eta_i,p_0)+\Delta\pi(\eta_i,H_R)-\kappa(\tau_i,p_0)\right]\Big|\beta^*\right].$$

Using the fact that

$$\begin{aligned}\mathbb{E}_\nu[\Delta\pi(\eta_i,H_R)|i\in\Lambda,\beta^*]&=\mathbb{E}_\nu\left[\sum_{\tau\in\eta_i}\left(\mu_\Lambda(\tau|\eta_i)-\mu_I(\tau|\eta_i)\right)\left(\kappa(\tau,p_0)-\widehat{\kappa}(\eta_i,p_0)\right)\Big|i\in\Lambda,\beta^*\right]\\&=\mathbb{E}_\nu\left[\sum_{\tau\in\eta_i}\left(\mu_\Lambda(\tau|\eta_i)-\mu_I(\tau|\eta_i)\right)\left(\kappa(\tau,p_0)-\kappa(\eta_i,p_0)\right)\Big|i\in\Lambda,\beta^*\right]\\&=\mathbb{E}_\nu\left[\sum_{\tau\in\eta_i}\mu_\Lambda(\tau|\eta_i)(\kappa(\tau,p_0)-\kappa(\eta_i,p_0))\Big|i\in\Lambda,\beta^*\right]\end{aligned}$$

and the fact that

$$\mathbb{E}_\nu\left[\sum_{i\in\Lambda}\kappa(\eta_i,p_0)-\kappa(\tau_i,p_0)+\Delta\pi(\eta_i,H_R)\right]=\mathbb{E}_\nu\left[\sum_{\eta\in E}|\Lambda^\eta|\left[\Delta\pi(\eta,H_R)+\sum_{\tau\in\eta}\mu_\Lambda(\tau|\eta)[\kappa(\eta,p_0)-\kappa(\tau,p_0)]\right]\right]$$

we obtain that indeed,

$$\mathbb{E}_\nu\left[\sum_{i\in I}\lambda(\tau_i)\left[\kappa(\eta_i,p_0)+\Delta\pi(\eta_i,H_R)-\kappa(\tau_i,p_0)\right]\Big|\beta^*\right]=0,$$

and hence, for any reporting strategy β, $U_1^B\leq 0$, which yields point (i).

Point (ii) follows from the fact that whenever λ is measurable with respect to E, then for all reporting strategies β

$$\mathbb{E}_\nu\left[\sum_{i\in I}\lambda(\tau_i)(\kappa(\eta_i,p_0)-\kappa(\tau_i,p_0))\right]=0$$

and

$$\mathbb{E}_\nu \left[\Delta\pi(\eta, H)\right] = \mathbb{E}_\nu \left[\sum_{\tau\in\eta} [\mu_\Lambda(\tau|\eta) - \mu_I(\tau|\eta)] r_\tau^H\right] = 0,$$

where $r_\tau^H = \frac{1}{|H^\tau|}\sum_{i\in H^\tau} \widehat{c}_i(p_0) - \widehat{\kappa}(\eta, p_0)$ denotes the mean residual of the baseline capitation formula computed in the hold-out sample.

Finally point (iii) follows from the fact that

$$\begin{aligned}
U_1^C &\leq \max_{\lambda\in[0,1]^T, E\in\mathcal{E}} \mathbb{E}_\nu \left[\sum_{i\in I} \lambda(\tau_i)(\widehat{\kappa}(\eta_i, p_0) - \kappa(\eta_i, p_0))\right] \\
&\leq \max_{\lambda\in\mathcal{M}(E), E\in\mathcal{E}} \mathbb{E}_\nu \left[\sum_{i\in I} \lambda(\tau_i)(\widehat{\kappa}(\eta_i, p_0) - \kappa(\eta_i, p_0))\right] \\
&\leq \max_{\lambda\in\mathcal{M}(E), E\in\mathcal{E}} \mathbb{E}_\nu \left[\sum_{\eta\in E}\left[\sum_{i\in I^\eta} \widehat{\kappa}(\eta_i, p_0) - \kappa(\eta_i, p_0)\right]^+\right],
\end{aligned}$$

which yields point (iii). $\square$

Proof of Lemma 1: We begin with point (i) and show that $\mathbb{E}_\nu\left[\Psi(\mathcal{E}, e)\right] \leq \mathbb{E}_\nu[\Psi(\mathcal{E}, e')]$ using a coupling argument, i.e. by carefully jointly sampling original errors e and Rademacher errors e'.

Consider the following process for generating errors e and e'. Errors e are generated according to the original distribution of e_i (where the different e_i's are independent of one another). In turn, each error term e'_i is generated from e_i as follows: conditional on e_i, $e'_i \in \{-c_{max}, c_{max}\}$ is chosen so that $\mathbb{E}_\nu[e'_i|e_i] = e_i$. Note that this is possible since $e_i \in [-c_{max}, c_{max}]$, and there is a unique such distribution. Since error terms $(e_i)_{i\in D_0}$ are independent, so are error terms $(e'_i)_{i\in I}$. In addition,

$$\mathbb{E}_\nu[e'_i] = \mathbb{E}_{e_i}\mathbb{E}_\nu[e'_i|e_i] = \mathbb{E}_{e_i} e_i = 0,$$

which implies $e'_i \sim U\{-c_{max}, c_{max}\}$.

We now show that necessarily $\mathbb{E}_\nu\left[\Psi(\mathcal{E},e)\right] \leq \mathbb{E}_\nu[\Psi(\mathcal{E},e')]$. Note that $\Psi(\mathcal{E},e)$ can be viewed as the maximum value for $S \subset E \in \mathcal{E}$ of

$$\Sigma_S \equiv \sum_{\eta\in S} |I^\eta| \left[\sum_{\tau\in\eta} \mu_I(\tau|\eta)\frac{1}{|D_0^\tau|}\sum_{i\in D_0^\tau} e_i\right].^{19} \tag{21}$$

Fix e, and assume that $\Psi(\mathcal{E},e)$ is realized by Σ_S for some set S of η's. We have by linearity of expectation that

$$\begin{aligned}\Psi(\mathcal{E},e) &= \sum_{\eta\in S} |I^\eta| \left[\sum_{\tau\in\eta} \frac{|I^\tau|}{|I^\eta||D_0^\tau|}\sum_{i\in D_0^\tau} e_i\right] = \mathbb{E}_\nu\left[\sum_{\eta\in S} |I^\eta| \left[\sum_{\tau\in\eta} \frac{|I^\tau|}{|I^\eta||D_0^\tau|}\sum_{i\in D_0^\tau} e'_i\right] \Big| e\right] \\ &\leq \mathbb{E}_\nu\left[\Psi(\mathcal{E},e')|e\right].\end{aligned}$$

Using the law of iterated expectations, this completes the proof of point (i).

We now turn to point (ii). Let $E \in \mathcal{E}$ be any partition, and let $S \subset E$ be a selection of elements in partition E. We first show that for all t,

$$\text{prob}\left(\Sigma_S > t\right) \leq \exp\left(-\frac{t^2|D_0|}{2c_{max}^2\alpha^2|I|^2}\right) \tag{22}$$

where Σ_S is defined by (21). Using Hoeffding's inequality (see Hoeffding (1963) or Cesa-Bianchi and Lugosi (2006), Lemma 2.2) we have that

$$\begin{aligned}\text{prob}\left(\Sigma_S > t\right) &= \exp\left[-\frac{2t^2}{\sum_{\eta\in S,\tau\in\eta}\sum_{i\in D_0^\tau} 4c_{\max}^2\frac{|I^\tau|^2}{|D_0^\tau|^2}}\right] \\ &\leq \exp\left[-\frac{t^2}{2c_{\max}^2\sum_{\tau\in T}\frac{|I^\tau|^2}{|D_0^\tau|}}\right] = \exp\left[-\frac{t^2}{2c_{\max}^2\frac{|I|^2}{|D_0|}\sum_{\tau\in T}\frac{|I^\tau|}{|D_0^\tau|}\frac{|D_0|}{|I|}\frac{|I^\tau|}{|I|}}\right] \\ &\leq \exp\left[-\frac{t^2}{2c_{\max}^2\frac{|I|^2}{|D_0|}\alpha}\right].\end{aligned}$$

[19]Indeed, the corresponding set S will only select ηs such that $\sum_{\tau\in\eta}\mu_I(\tau|\eta)\frac{1}{|D_0^\tau|}\sum_{i\in D_0^\tau} e_i > 0$.

Since there are at most $M = \sum_{E\in\mathcal{E}} 2^{|E|} - 1$ possible non-empty sets S, this implies that

$$\text{prob}\left(\Psi(\mathcal{E}, e) > t\right) \leq M \exp\left[-\frac{t^2}{2c_{\max}^2 \frac{|I|^2}{|D_0|}\alpha}\right].$$

To complete the proof, we use the fact that $\mathbb{E}_\nu[\Psi(\mathcal{E}, e)] = \int_0^{+\infty} \text{prob}(\Psi(\mathcal{E}, e) > t)\mathrm{d}t$. Pick t_0 such that $M \exp\left[-\frac{t_0^2}{2c_{\max}^2 \frac{|I|^2}{|D_0|}\alpha}\right] = 1$, i.e. $t_0 = |I| c_{\max}\sqrt{\frac{2\alpha \log M}{|D_0|}}$. We have

$$\begin{aligned}
\mathbb{E}_\nu[\Psi(\mathcal{E}, e)] &\leq \int_0^{t_0} \text{prob}(\Psi(\mathcal{E}, e) > t)\mathrm{d}t + \int_{t_0}^{+\infty} \text{prob}(\Psi(\mathcal{E}, e) > t)\mathrm{d}t \\
&\leq t_0 + \int_{t_0}^{+\infty} M \exp\left[-\frac{t^2}{2c_{\max}^2 \frac{|I|^2}{|D_0|}\alpha}\right] \\
&\leq |I| c_{\max}\sqrt{\frac{2\alpha \log M}{|D_0|}} + \frac{\sqrt{2\pi}}{2}|I| c_{\max}\sqrt{\frac{\alpha}{|D_0|}} M \exp\left[-\frac{t^2}{2c_{\max}^2 \frac{|I|^2}{|D_0|}\alpha}\right] \\
&\leq |I| c_{\max}\sqrt{\frac{2\alpha}{|D_0|}}\left(\sqrt{\log M} + 1\right).
\end{aligned}$$

□

Proof of Proposition 5: Let E be the partition maximizing $\sum_{\eta\in E} \frac{1}{1+\sqrt{|D_0^\eta|}}\mu_I(\eta)$. We start with the following simple claim:

Claim 1 (hard to distinguish distributions)**.** *For each integer $d \geq 0$, there exists a pair of distributions ϕ_0, ϕ_1 with finite support over $[0, c_{max}]$ such that $\mathbb{E}_{\phi_0} c = c_l^d$, $\mathbb{E}_{\phi_1} c = c_h^d$, $c_h^d, c_l^d \in [c_{max}/4, 3c_{max}/4]$, $c_h^d - c_l^d \geq k' c_{max}/(1 + \sqrt{d})$, and ϕ_0^d is hard to distinguish from ϕ_1^d, in the sense that*

$$\sup_{S\subset[0,c_{\max}]^d} \phi_0^d(S) - \phi_1^d(S) \leq 1/4,$$

for some universal constant $k' > 0$, where ϕ_0^d and ϕ_1^d denote the d product measures.

We defer the proof of Claim 1 until after the proof of the proposition. We use the notation $d(\eta) \equiv |D_\eta^0|$. It is sufficient for our lower bound to consider the following class of environments ν.

- Selection cost $K(\cdot)$ is identically equal 0.
- Cost distributions for the public and private plans are determined as follows. Let $(b_\eta)_{\eta \in E}$ be independent Bernoulli draws such that $\text{prob}(b_\eta = 1) = 1/2$. For all $\tau \in \eta$, cost distributions $c(p_0, \tau)$ are independent and identically distributed according to the distribution $\phi_{b_\eta}^{d(\eta)}$ described in Claim 1. Its expected value is $c_0^\eta \in \{c_h^{d(\eta)}, c_l^{d(\eta)}\}$.
- For all $\tau \in \eta$, the private plans' cost $c(p_1, \tau_i)$ is distributed according to $\frac{1}{3}(\phi_h^{d(\eta)} + \phi_l^{d(\eta)} + c(p_0, \tau))$.
- Holdout set H contains sufficient information to identify $(b_\eta)_{\eta \in E}$.
- Private plan p_1 knows $(b_\eta)_{\eta \in E}$.

For notational convenience, we denote by $\underline{c}(p_j, \eta)$ and $\overline{c}(p_j, \eta)$ the cost distributions for plan j and characteristic η when b_η is respectively equal to 0 and 1. More generally, denote by $\underline{c}(p_j)$ the vector of expected per-patient cost functions for p_j assuming $b_\eta = 0$. Note that $\underline{c}(p_j)$ and $\overline{c}(p_j)$ will agree on patients outside of η.

Let $g \in \mathcal{G}_\nu$ be an incentive compatible direct-revelation mechanism. Fix an $\eta \in E$, and a realization of D_0^η. We derive a lower bound for the efficiency loss incurred by g over patients with characteristic η (the number of such patients is $|I| \cdot \mu_I^E(\eta)$).

We exploit incentive compatibility conditions using the following set of messages. Message $\overline{m}_0$ is the message of public plan p_0 that correctly reports $(b_{\eta'})_{\eta' \neq \eta}$ but reports $b_\eta = 1$. Messages $\underline{m}_0$, $\overline{m}_1$, $\underline{m}_1$ are defined similarly. Note that message m_1 affects both transfers $\Pi(\mathcal{D}_0, m_0, m_1)$ and the selection of patients $\lambda(m_1)$.

For notational convenience, we will treat distribution $\lambda(\cdot)$ as a vector. Throughout, we take expectations over the realization of $b_{-\eta}$ and $\widehat{c}$ (cost indicators for groups other than η, and realized costs of care). Thus, for example, $\mathbb{E}_{b_{-\eta}, \widehat{c}} \langle \overline{c}(p_1), \lambda(\overline{m}_1) \rangle$ is the expected cost of care for private plan p_1 assuming that $b_\eta = 1$; and $\mathbb{E}_{b_{-\eta}, \widehat{c}} \langle \overline{c}(p_1), \lambda(\underline{m}_1) \rangle$ is the expected cost accrued to p_1 from treating its patients when $b_\eta = 1$, but p_1 reports that $b_\eta = 0$. We drop the $b_{-\eta}, \widehat{c}$ subscript from now on.

Incentive compatibility of plan p_1's messages if $b_\eta = 1$ implies that

$$\mathbb{E}_\nu \Pi(\mathcal{D}_0, \overline{m}_0, \overline{m}_1) - \mathbb{E}_\nu \langle \overline{c}(p_1), \lambda(\overline{m}_1)\rangle \geq \mathbb{E}_\nu \Pi(\mathcal{D}_0, \overline{m}_0, \underline{m}_1) - \mathbb{E}_\nu \langle \overline{c}(p_1), \lambda(\underline{m}_1)\rangle. \tag{23}$$

Incentive compatibility of plan p_0's message when $b_\eta = 0$ implies that

$$-\mathbb{E}_\nu \Pi(\mathcal{D}_0, \underline{m}_0, \underline{m}_1) + \mathbb{E}_\nu \langle \underline{c}(p_0), \lambda(\underline{m}_1)\rangle \geq -\mathbb{E}_\nu \Pi(\mathcal{D}_0, \overline{m}_0, \underline{m}_1) + \mathbb{E}_\nu \langle \underline{c}(p_0), \lambda(\underline{m}_1)\rangle,$$

which simplifies to

$$\mathbb{E}_\nu \Pi(\mathcal{D}_0, \underline{m}_0, \underline{m}_1) \leq \mathbb{E}_\nu \Pi(\mathcal{D}_0, \overline{m}_0, \underline{m}_1). \tag{24}$$

Combining (23) and (24) we obtain that

$$\mathbb{E}_\nu \Pi(\mathcal{D}_0, \overline{m}_0, \overline{m}_1) - \mathbb{E}_\nu \Pi(\mathcal{D}_0, \underline{m}_0, \underline{m}_1) \geq \mathbb{E}_\nu \langle \overline{c}(p_1), \lambda(\overline{m}_1) - \lambda(\underline{m}_1)\rangle. \tag{25}$$

A symmetric argument implies that

$$\mathbb{E}_\nu \Pi(\mathcal{D}_0, \overline{m}_0, \overline{m}_1) - \mathbb{E}_\nu \Pi(\mathcal{D}_0, \underline{m}_0, \underline{m}_1) \leq \mathbb{E}_\nu \langle \underline{c}(p_1), \lambda(\overline{m}_1) - \lambda(\underline{m}_1)\rangle. \tag{26}$$

Together, (25) and (26) imply

$$\mathbb{E}\langle \overline{c}(p_1) - \underline{c}(p_1), \lambda(\overline{m}_1) - \lambda(\underline{m}_1)\rangle \leq 0. \tag{27}$$

Since $\overline{c}(p_1) - \underline{c}(p_1)$ is a positive constant on η and 0 elsewhere, (27) implies that in expectation at least as many patients from η are treated by p_1 when $b_\eta = 0$ as when $b_\eta = 1$. Note that the efficiency loss that occurs when a patient $i \in \eta$ is treated by p_0 when $b_\eta = 1$ or is treated by p_1 when $b_\eta = 0$ is $(c_h^{d(\eta)} - c_l^{d(\eta)})/3$. Denote by L_0 the expected loss per patient in η if

$b_\eta = 0$, and by L_1 the expected loss per patient if $b_\eta = 1$. We thus have

$$\text{prob}[p_1 \text{ treats}|b_\eta = 0] \geq \text{prob}[p_1 \text{ treats}|b_\eta = 1] = 1 - \text{prob}[p_0 \text{ treats}|b_\eta = 1],$$

and

$$L_0+L_1 = ((c_h^{d(\eta)} - c_l^{d(\eta)})/3)\cdot(\text{prob}[p_1 \text{ treats}|b_\eta = 0]+\text{prob}[p_0 \text{ treats}|b_\eta = 1]) \geq (c_h^{d(\eta)} - c_l^{d(\eta)})/3. \tag{28}$$

Define $q_\eta \equiv \text{prob}[b_\eta = 1|\mathcal{D}_0]$. The expected efficiency loss accrued per patient in η is greater than

$$\min(L_0, L_1)\cdot\max(q_\eta, 1-q_\eta) + \max(L_0, L_1)\cdot\min(q_\eta, 1-q_\eta) \geq \\ (L_0 + L_1)\cdot\min(q_\eta, 1-q_\eta) \geq (c_h^{d(\eta)} - c_l^{d(\eta)})\cdot\min(q_\eta, 1-q_\eta)/3. \tag{29}$$

Exploiting Claim 1 we will show that

$$\mathbb{E}\min(q_\eta, 1-q_\eta) = \mathbb{E}_{D_0^\eta}\min\left(\frac{\phi_0(D_0^\eta)}{(\phi_0+\phi_1)(D_0^\eta)}, \frac{\phi_1(D_0^\eta)}{(\phi_0+\phi_1)(D_0^\eta)}\right) \geq \frac{1}{4}, \tag{30}$$

where we abuse notation and set $\phi_0(D_0^\eta) = \prod_{\widehat{c}_i\in D_\eta^0}\phi_0(\widehat{c}_i)$. Note that D_0^η is distributed according to $(\phi_0^{d(\eta)} + \phi_1^{d(\eta)})/2$. The first equality of (30) holds by Bayes rule, and the fact that b_η is a uniform Bernoulli. Furthermore, Claim 1 implies that

$$\mathbb{E}_{D_0^\eta}\min\left(\frac{\phi_0(D_0^\eta)}{(\phi_0+\phi_1)(D_0^\eta)}, \frac{\phi_1(D_0^\eta)}{(\phi_0+\phi_1)(D_0^\eta)}\right) = \\ \mathbb{E}_{D_0^\eta}\left[\frac{1}{2} - \frac{|\phi_0(D_0^\eta) - \phi_1(D_0^\eta)|}{2(\phi_0+\phi_1)(D_0^\eta)}\right] = \frac{1}{2} - \sum_{D_0^\eta}\frac{|\phi_0(D_0^\eta) - \phi_1(D_0^\eta)|}{4} > \frac{1}{4}.$$

Combining (29) and (30) it follows that per-patient efficiency loss in η is at least

$$(c_h^{d(\eta)} - c_l^{d(\eta)}) \times \min(q_\eta, 1-q_\eta)/3 \geq (c_h^{d(\eta)} - c_l^{d(\eta)})/12 = \frac{(k'/12)c_{max}}{1+\sqrt{|D_0^\eta|}}.$$

Setting $k = k'/12$ completes the proof. $\square$

We now prove Claim 1.

Proof of Claim 1: Given $d \geq 1$, let $\phi_0 \sim c_{\max} B_{1/2-\varepsilon}$, and $\phi_1 \sim c_{max} \cdot B_{1/2+\varepsilon}$, where B_q denotes Bernoulli variables of parameter q, and $0 < \varepsilon < 1/4$ (with the relationship between ε and d to be specified below). Standard results from information theory (Cover and Thomas, 2012) imply that the statistical distance between ϕ_0^d and ϕ_1^d satisfies

$$2 \sup_{S \subset [0, c_{\max}]^d} \phi_0^d(S) - \phi_1^d(S) \leq \sqrt{d \cdot D(B_{1/2+\varepsilon} \| B_{1/2})/2} = \sqrt{d \cdot O(\varepsilon^2)} < k_1 \cdot \varepsilon\sqrt{d},$$

where $D(\cdot\|\cdot)$ is the Kullback-Leibler divergence, and $k_1 \geq 2$ is a constant. Choose $\varepsilon = 1/(2k_1\sqrt{d}) \leq 1/4$. Claim 1 holds with

$$c_h^d - c_l^d = 2\varepsilon = 1/(k_1\sqrt{d}).$$

Setting $k' \leq 1/k_1$ completes the proof. $\square$

B.3 Proofs for Section 5

Proof of Proposition 6: Consider the subgame following entry. For any continuation pricing equilibrium (π_{p_0}, π_{p_1}), the usual Bertrand competition argument implies that price formulas must satisfy

$$\forall \eta, \pi_0(\eta) = \pi_1(\eta) = \kappa(\eta).$$

Given these prices, profits are determined by the plans' selection behavior $(\lambda_{p_0}, \lambda_{p_1})$. Given the selection rule $\lambda_{\neg p}$ of her competitor, plan p chooses

$$\begin{aligned}\lambda_p &\in \arg\max_{\lambda\in[0,1]^T} \sum_{\tau\in T}(\kappa(\eta)-\kappa(\tau))\left(\frac{\mu_I}{2}(\tau)+\lambda(\tau)-\lambda_{\neg p}(\tau)\right)-K(\lambda)\\ &= \arg\max_{\lambda\in[0,1]^T} \sum_{\tau\in T}\lambda(\tau)(\kappa(\eta)-\kappa(\tau))-K(\lambda).\end{aligned}$$

Since K is strictly convex, minimized at $\mu_I/2$, and smooth, it follows that its gradient $\nabla K_{|\mu_I/2}$ at $\mu_I/2$ is equal to 0. As a result both plans engage in the same non-zero amount of selection λ^*, so that in aggregate, selection has no effect on each plan's treated sample. Strict convexity of K implies that $K(\lambda^*) > 0$. This means that the entrant gets strictly negative expected profits following entry.

It follows that the unique equilibrium involves no entry, allowing the incumbent to charge prices equal to $\overline{\pi}$. □

Proof of Proposition 7: Consider the subgame following entry. For any continuation pricing equilibrium (π_{p_0}, π_{p_1}), the usual Bertrand competition argument implies that price formulas must satisfy

$$\forall\eta, \pi_0(\eta)=\pi_1(\eta)=\kappa(\eta).$$

Assuming truthful reporting by plan $\neg p$, strategic capitation ensures that plan p does not benefit from selecting a non representative sample of types. Hence plan p's payoffs boils down to

$$\sum_{\tau\in T}(\kappa(\eta)-\kappa(\tau))\left(\mu_I(\tau)-\lambda_{\neg p}(\tau)\right)-K(\lambda).$$

It is therefore optimal for plan p to set $\lambda_p = \mu_I/2$ and minimize selection cost. Given this choice, it is indeed optimal for plan $\neg p$ to report its hold-out sample truthfully. □

References

Antic, N. (2014): "Contracting with Unknown Technologies," *unpublished paper, Princeton University.*

Batata, A. (2004): "The effect of HMOs on fee-for-service health care expenditures: evidence from Medicare revisited," *Journal of health economics*, 23, 951–963.

Belloni, A., V. Chernozhukov, I. Fernández-Val, and C. Hansen (2013): "Program evaluation with high-dimensional data," *arXiv preprint arXiv:1311.2645.*

Belloni, A., V. Chernozhukov, and C. Hansen (2014): "Inference on Treatment Effects after Selection among High-Dimensional Controls," *The Review of Economic Studies*, 81, 608–650.

Bergemann, D. and K. H. Schlag (2008): "Pricing without priors," *Journal of the European Economic Association*, 6, 560–569.

Bisin, A. and P. Gottardi (1999): "Competitive equilibria with asymmetric information," *Journal of Economic Theory*, 87, 1–48.

——— (2006): "Efficient competitive equilibria with adverse selection," *Journal of political Economy*, 114, 485–516.

Brooks, B. (2014): "Surveying and selling: Belief and surplus extraction in auctions," .

Brown, J., M. Duggan, I. Kuziemko, and W. Woolston (2014): "How Does Risk Selection Respond to Risk Adjustment? New Evidence from the Medicare Advantage Program," *American Economic Review*, 104, 3335–3364.

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

Carroll, G. (2013): "Robustness and Linear Contracts," *Stanford University Working Paper.*

——— (2015): "Robustness and Separation in Multidimensional Screening," *Stanford University Working Paper.*

Cesa-Bianchi, N. and G. Lugosi (2006): *Prediction, Learning, and Games*, Cambridge University Press.

Chassang, S. (2013): "Calibrated incentive contracts," *Econometrica*, 81, 1935–1971.

Cover, T. M. and J. A. Thomas (2012): *Elements of information theory*, John Wiley & Sons.

Cremer, J. and R. P. McLean (1988): "Full extraction of the surplus in Bayesian and dominant strategy auctions," *Econometrica: Journal of the Econometric Society*, 1247–1257.

Dubey, P. and J. Geanakoplos (2002): "Competitive pooling: Rothschild-stiglitz reconsidered," *Quarterly Journal of Economics*, 1529–1570.

Ellis, R. P. (1998): "Creaming, skimping and dumping: provider competition on the intensive and extensive margins," *Journal of health economics*, 17, 537–555.

Epple, D., D. Figlio, and R. Romano (2004): "Competition between private and public schools: testing stratification and pricing predictions," *Journal of Public Economics*, 88, 1215–1245.

Evans, M. A., G. C. Pope, J. Kautter, M. J. Ingber, S. Freeman, R. Sekar, and C. Newhart (2011): "Evaluation of the CMS-HCC Risk Adjustment Model," *CfMM Services, Editor.*

for Medicare, C., C. f. C. I. Medicaid Services, and I. Oversight (2016): "March 31, 2016, HHS-operated risk adjustment methodology meeting: discussion paper," .

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

Glazer, J. and T. G. McGuire (2000): "Optimal risk adjustment in markets with adverse selection: an application to managed care," *American Economic Review*, 1055–1071.

Hartline, J. D. and T. Roughgarden (2008): "Optimal Mechanism Design and Money Burning," in *Symposium on Theory Of Computing (STOC)*, 75–84.

Hoeffding, W. (1963): "Probability Inequalities for Sums of Bounded Random Variables," *Journal of the American Statistical Association*, 58, 13–30.

Kautter, J., G. C. Pope, M. Ingber, S. Freeman, L. Patterson, M. Cohen, and P. Keenan (2014): "The HHS-HCC risk adjustment model for individual and small group markets under the Affordable Care Act," *Medicare & Medicaid research review*, 4.

Madarász, K. and A. Prat (2014): "Screening with an Approximate Type Space," *Working Paper, London School of Economics.*

Massart, P. and J. Picard (2007): *Concentration inequalities and model selection*, vol. 1896, Springer.

Mello, M. M., S. C. Stearns, E. C. Norton, and T. C. Ricketts (2003): "Understanding biased selection in Medicare HMOs," *Health services research*, 38, 961–992.

Mezzetti, C. (2004): "Mechanism design with interdependent valuations: Efficiency," *Econometrica*, 72, 1617–1626.

Newhouse, J. P., M. Price, J. Huang, J. M. McWilliams, and J. Hsu (2012): "Steps to reduce favorable risk selection in Medicare advantage largely succeeded, boding well for health insurance exchanges," *Health Affairs*, 31, 2618–2628.

Pope, G. C., J. Kautter, R. P. Ellis, A. S. Ash, J. Z. Ayanian, M. J. Ingber, J. M. Levy, and J. Robst (2004): "Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Model," *HEALTH CARE FINANCING REVIEW*, 25, 119.

Riordan, M. H. and D. E. Sappington (1988): "Optimal contracts with public ex post information," *Journal of Economic Theory*, 45, 189–199.

Roth, A. E. (2007): "Repugnance as a Constraint on Markets," *The Journal of Economic Perspectives*, 21, 37–58.

Rothschild, M. and J. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *The Quarterly Journal of Economics*, 90, 629–649.

Segal, I. (2003): "Optimal pricing mechanisms with unknown demand," *The American economic review*, 93, 509–529.

Starc, A. (2014): "Insurer pricing and consumer welfare: Evidence from medigap," *The RAND Journal of Economics*, 45, 198–220.

Vapnik, V. (1998): *Statistical learning theory*, Wiley New York.

Walters, C. (2012): "A Structural Model of Charter School Choice and Academic Achievement," *working paper, University of California, Berkeley.*