Inverse selection*

Markus K. Brunnermeier Rohit Lamba Carlos Segura-Rodriguez

November 2023

Abstract

Big data and AI inverts adverse selection problems. It allows insurers to infer statistical information and thereby reverses information advantage from the insuree to the insurer. In a setting with two-dimensional type space whose correlation can be inferred with big data we derive three results: First, a novel trade-off between obfuscation and price discrimination—the insurer tries to exploit but also protect its statistical information by offering few screening contracts. Second, insuree's ability to do perfect Bayesian inference limits the returns to inverse selection for the insurer. Third, competition and forced transparency reduces total surplus and insurer's payoff while increasing insuree's payoff.

Keywords: Insurance, Big Data, Informed Principal, Obfuscation, Price Discrimination.

JEL Classification: G22, D82, D86, C55.

*Brunnermeier: markus@princeton.edu, Princeton University; Lamba: rlamba@psu.edu, Pennsylvania State University; Segura-Rodriguez: segurarc@bccr.fi.cr Banco Central de Costa Rica. We are indebted to Nemanja Antić, Tommaso Denti, Goutham Gopalakrishna, Hanming Fang (discussant), Ben Handel, Emir Kamenica, Victoria Marone, Stephen Morris, Rakesh Vohra, Motohiro Yogo, and participants at various seminars and conferences for useful comments.

1 Introduction

With the rise of big data, artificial intelligence (AI), and machine learning almost every action is recorded, and correlates are constructed to better predict customers' behavior. The insurance industry is no exception and is undergoing a radical transformation — contract design and price discrimination will fundamentally change.[1]

In most information models, customers have an informational advantage and the principal, e.g. the insurance company, faces an adverse selection problem. It tries to mitigate this information disadvantage by offering a menu of screening contracts to potential customers.[2] Big data and new statistical tools allow insurance companies to infer superior correlates about customer characteristics – over which customers still have private information. These correlates enable the insurer to obtain informational advantage about the ultimate risk. In other words, the principal can "invert" the mapping from characteristics to risks. Thus, big data and AI transform many adverse selection problems into what we call "inverse selection" problems.

The objective of this paper is fourfold: First, to introduce a principal-agent model that captures the aforementioned greater statistical informational advantage for the principal. Second, to introduce related economic taxonomy that arises due to new trade-offs that emerge from inverse selection. Third, to show under what modeling assumptions the returns to statistical informational advantage are high for the principal; it is the principal's ability to choose information disclosure and contract design policies at the outset, and the agent's limited ability to do Bayesian inference. And, fourth, to connect the predictions of the model to market realities such as the rise of data brokers and importance of consumer activism.

Our setting is close in spirit to the informed principal approach in mechanism design (Myerson [1983] and Maskin and Tirole [1990, 1992]). It departs from the canonical structure in two ways: First, while the agent has hard private information— family history, eating habits, zip code, etc; the principal has statistical private information— how these characteristics interact and determine the agent's probability of, say, getting cancer. Second, there is an information design problem of how the principal uses the statistical information layered on top of a mechanism design problem of choosing contracts. Overall, the basic structure here is inspired from the classical insurance model studied by Rothschild and Stiglitz [1976] with two key differences: we consider a richer information structure, and for the most part restrict attention to a monopolistic screening setup.

Inverse selection not only differ from standard *adverse selection* but also from the more recent *advantageous selection* literature. Advantageous selection stresses the importance of preference het-

[1]For instance, the International 'Geneva' Association for the Study of Insurance Economics states in a recent report that "Advances in big data analytics, artificial intelligence and the Internet of Things promise to fundamentally transform the insurance industry and the role data plays in insurance. New sources of digital data, for example in online media and the Internet of Things, reveal information about behaviours, habits and lifestyles that allows us to assess individual risks much better than before."— Keller et al. [2018].

[2]Akerlof [1970] pioneered the study of adverse selection and screening. The core idea has found applications is variety of settings: Rothschild and Stiglitz [1976] study the insurance problem, Mailath and Postlewaite [1990] study public goods provision, and Biais, Martimort, and Rochet [2000] and Tirole [2012] study various aspects of financial markets, to name a few. See Green and Laffont [1979] and Laffont and Martimort [2009] for general theoretical treatments of the principal agent screening problem.

erogeneity in order to overturn the standard theoretical, but empirically counterfactual, result that the high-risk agents get full insurance whereas low-risk agents opt for partial insurance. With preference heterogeneity, highly risk-averse agents buy more insurance, despite the fact that they are less risky, since they behave more cautiously.[3] In both settings, adverse and advantageous selection, the insurance provider suffers from an informational disadvantage. This is in contrast to our inverse selection setting in which private statistical information is held by the insurance company.

We model the inverse selection problem using a two-dimensional type space. Both dimensions determine the riskiness of the agent: The agent perfectly knows one (type of) characteristic, the first dimension of the type. In contrast, the principal, e.g. the insurer, privately knows the correlation across both dimensions. The marginal distribution along both dimensions, though, is common knowledge. At a high level, we equip the agent with greater hard or physical information and the principal with greater soft or statistical information.[4]

The basic tension the principal faces is the following: She has to decide the sensitivity of the set of offered contracts to the statistical correlation. The more responsive the contract is to the underlying correlation, the greater price discrimination can be achieved. However, in doing so, she has to be aware that by offering more fine-tuned screening contracts, she may partially reveal her (statistical) informational advantage. In other words, the principal faces an *obfuscation-versus-price discrimination* trade-off. By offering a richer set of contracts, the principal can screen and discriminate more but will also end up giving up some of its statistical informational advantage, which the agent can then use while picking the contract. Note that this trade-off is different from *rent-versus-efficiency* prevalent in standard principal agent problems, where the principal worsens efficient risk-sharing in order to minimize the information rent that the agent can extract. Of course, the standard rent-versus-efficiency trade-off is also present in our setting with respect to the agent's private information.

As in the classical setup, the optimal contract separates along the insuree's private information. However, along the private statistical information of the insurer, the optimal contract features some pooling (possibly complete pooling); we show that complete separation along both dimensions is never optimal for the insurer. When the insurer pools certain correlation types, she is giving up on price discrimination in order to maintain the statistical information advantage.

Our formal analysis builds on the information and contract design problem, wherein the principal can fully specify at the outset the information structure that informs the agent about the statistical correlation, and on the basis of this disclosure, an insurance contract is offered. In terms of techniques, the information design part follows Kamenica and Gentzkow [2011] and Bergemann and Morris [2016], and the contract design problem takes this as given in solving for a standard screening contract (as in Börgers [2015], Chapter 2).

The principal can always choose full disclosure, in which case the model collapses to the Roth-

[3]Einav and Finkelstein [2011] provide an overview of the key ideas. See Finkelstein and McGarry [2006] and Fang, Keane, and Silverman [2008] for empirical evidence on adverse and advantageous selection.

[4]The model can be equivalently interpreted as the first dimension being the set of all characteristics and the second dimension being the riskiness of the agent. Then the agent has private information about personal characteristics, and the principal understands the mapping between characteristics and risks.

schild and Stiglitz [1976] benchmark.[5] This is, of course, not optimal as the principal can do better. In fact, the optimal information disclosure strikingly assumes a bang-bang structure: two regions, one with complete pooling and the other with full information disclosure of the principal's information. The complete pooling regions shifts all mass to the unique correlation type, which makes the agent's information worthless so the principal offers full insurance and extracts full surplus, illustrating the power of inverse selection. In the complementary region, information is fully disclosed and the principal attains the RS profit. So overall, the principal does weakly better, point-wise, than the benchmark RS model and strictly better in the pooling region.

Next, we explore the implications of inverse selection when the principal decides on the information structure ex-post, and hence the model reduces to a cheap-talk like first stage (Crawford and Sobel [1982]) followed by a contract design problem. In such a scenario, revealing small bits of information is too costly, so that the insurer prefers to offer only a finite number of contracts.[6] In the language of Myerson [1981], the optimum features *ironing* almost everywhere. In fact, we further show that the number of contracts turns out to be small, highlighting that the obfuscation-price discrimination trade-off is firmly resolved in the favor of the former. In most cases, the contract space along the statistical information is partitioned into one or two contracts. In fact, for a large class of parameters, when the principal offers only one contract, she prefers not to use her information in contract design at all. Thus, when the insurer cannot commit to the information disclosure policy, the potential gains from inverse selection can be significantly curtailed.

Importantly, we show that it is the agents' ability to rationally infer part of the principal's information from her contract offering that limits the principal's inverse selection profit. Perfect Bayesian inference when deciphering insurance contracts is, however, a tall order (see Handel and Kolstad [2015]). The agent may not only lack probabilistic sophistication (see Benjamin [2019]), but also the ability to invert the mapping from principal's contract offerings to her information advantage. To unpack this area of behavioral inverse selection, we look at two deviations from the standard Bayesian model.

First, we study the case in which the insuree is *gullible*, i.e., he believes whatever the insurer tells him about the correlation coefficient and does not infer any statistical information from the menu of contracts offered by the insurer. This model, although theoretically non-standard, clarifies the direction in which the insurer would like to push the contract if she could create the maximal obfuscation and implement the maximal price discrimination. The insurer misleads the insuree to the maximum extent possible, so only two correlations are ever reported— the lowest and highest possible values— and a distinct contract is chosen for each possible actual realization of the correla-

[5] When the statistical correlation is common knowledge, the problem collapses to the standard monopolistic Rothschild and Stiglitz [1976] insurance problem, but with a twist: which is the high type and which is low type in the agent's private information from the screening perspective depends on the realization of the statistical correlation. There is a unique correlation which splits which is the high or low type and exactly at that correlation the agent has no informational advantage.

[6] Eilat, Eliaz, and Mu [2020] study a standard quasi-linear monopolistic screening where the information change of the principal is exogenously restricted by a cap on KL-divergence between the prior and posterior. They too find that the number of contracts offered at the optimum is finite. Their model, mechanism and the application are however quite different than ours.

tion. The low-risk agent is overinsured, and the high-risk agent is often excluded from the market. The profits of the insurer are uniformly higher in comparison to the standard model.

Second, we say that the insuree is *naive* if he again does not infer any statistical information from the contract, but unlike gullible, sticks to the prior. Here too, the insurer gains on average, but ex post the ranking is not uniform: dictated by feasibility constraints, the insurer would like the insuree to update his belief (even correctly) in certain situations. For the naive case, the difference between the insurer's and insuree's beliefs is exogenously fixed by the prior, and the insurer maximizes on the price discrimination channel, given this constraint.[7]

Finally, we explore two extensions: First, we introduce competition into the model by allowing other 'regular' insurers to offer contracts who do not have access to the big data technology. They offer a single Rothschild-Stiglitz contract averaging over all possible correlations in order to screen the insuree along the first dimension. We show that competition sucks out some of the informational advantage of the insurer and increases the surplus of the agent. Second, the insurer is forced to put the statistical information (its bid data advantage) in the public domain. This increases consumer's profit but lowers insurer's profit. This increase is smaller than the decline, so that total surplus is also reduced.

While our model is admittedly stylistic, it provides a framework for policy makers on how to think about the role of big data and AI in the design of screening contracts in several ways. First, the contrast between our standard model and the gullible case shows that the returns to statistical information for the principal can be quite large, especially when the agents are not sophisticated. This points towards a market for acquiring consumer information, which in reality has manifested in the rise of data brokers such as Oracle, Nielsen and Salesforce.[8] Second, forcing the principals to make private statistical information public points towards the merits of a public data repository.[9] In fact, consumer activism and making them aware of how data is utilized may be a more effective way to limit exploitation through inverse selection than forcing the data to be public.[10] Finally, competition is another way to protect the consumer as it reduces the extent of price discrimination by big data monopolies.[11]

In addition to the aforementioned papers, our analysis is also related to several important strands of literature. Kamenica, Mullainathan, and Thaler [2011] explored early on the consequences of a firm knowing more about consumers' expected usage of the product for sale than the consumers themselves. This can be seen as a predecessor to what we call inverse selection. The informed principal problem seems to us as a likely candidate to make theoretical progress on the topic. To that end and to our knowledge, Villeneuve [2005] is the first paper to study insurance markets in the realm of the informed principal model. This has been followed up by Abrardi,

[7] In recent work, Fang and Wu [2020] also study a model of belief based behavioral biases and how these can be exploited systematically by firms in the market for insurance contracts.

[8] See for example, Financial Times [2019].

[9] See, for example, Rajan [2019].

[10] The call for transparency by the Federal Trade Commission (Ramirez et al. [2014]) and the framework for a general data protection regulation issued by the European Parliament (Council of the European Union [2016]) can be evaluated in this light.

[11] See, for example, Khan [2017].

Colombo, and Tedesch [2022], simultaneously, with our work. Both of these papers though focus on competing principals, in contrast to our monopolistic setup. Moreover, Villeneuve [2005], and for the most part, Abrardi et al. [2022] focus on one-dimensional private information on the side of the principal, whereas, we look at a two-dimensional state, part of which is known to the principal and part is known to the agent.[12]

In recent work, Luz, Gottardi, and Moreira [2023] and Bhaskar, McClellan, and Sadler [2023] also look at a two-dimensional type space for insurance contracts. The former considers heterogeneity in preferences, specifically risk-aversion, as the second dimension and the latter assumes that the first dimension is commonly known and can be used by a third party such as a regulator to offer a large number of contracts to implement the efficient outcome. The former focuses on foundations of advantageous selection and the latter on the regulation of insurance markets. Strausz [2023] studies the extent of price discrimination that can be achieved by a multi-product seller who can construct correlations in buyers' valuations using big data. While the setups and results of all these papers are quite different from ours, we view these papers as being complementary to our work in a push towards a deeper understanding of insurance models with the advent of modern data technology.

2 Setup

2.1 Primitives

The concept of inverse selection can be presented with a model of the informed principal (Myerson [1983]) as follows: A risk-neutral principal holds some statistical private information about an underlying state and can post a menu of insurance contracts to screen the agent. The agent is risk averse and holds some hard private information about the risk he faces.

Preferences. A profit maximizing monopolist insurer (principal/seller) interacts with an insuree (agent/buyer) who wants to insure himself against some damage/loss. The insurer is risk neutral and offers a standard insurance contract (p, x), where p represents the price (or premium), and x represents the proportion of the insuree's loss that is covered by the contract. So, $x < 1$ means under-insurance, $x = 1$ means full-insurance, and $x > 1$ means over-insurance.

The insuree has an initial wealth w. The uncertain loss he faces is a random variable with a well defined mean and variance. So, given a contract (p, x) and realized loss ℓ, his final wealth is given by $z = w - p - (1 - x)\ell$. The insuree is assumed to have a standard mean-variance preference. Thus, his utility is given by:

$$u(x, p) = \mathbb{E}[z] - \frac{\gamma}{2}\mathbb{V}[z] = w - p - (1 - x)\mu - \frac{\eta}{2}(1 - x)^2$$

where $\mu = \mathbb{E}[\ell]$ is the expected loss, $\mathbb{V}[\ell]$ the variance of loss, γ measures the extent of risk

[12]Beyond insurance markets, see also Mylovanov and Tröger [2014] and Koessler and Skreta [2019] for related theoretical models of the informed principal.

aversion, and $\eta = \gamma \times \mathbb{V}[\ell]$ captures the level of risk faced by the insuree. This expression can be simplified further as follows:

$$\begin{aligned} u(x,p) &= \underbrace{w-\mu}_{\varphi} + \underbrace{\left[x\mu - \frac{\eta}{2}(1-x)^2\right]}_{v(x)} - p \\ &= \varphi + v(x) - p. \end{aligned}$$

which lends a tractable structure to the insuree's preferences so that his utility is linear in money and concave in the extent of loss.[13]

Information. A standard approach to the insurance model would assume that the mean loss, μ, is the agent's private information. We depart from this crucial assumption on the "endowment" of information as follows. A relevant two-dimensional state $\theta = (\theta_1, \theta_2)$ determines μ, where $\theta_i \in \{a, b\}$ for $i \in \{1, 2\}$. So, given state θ, the mean loss of the agent is given by μ_θ. Without loss of generality, we assume that

$$\mu_{bb} > \mu_{ba} > \mu_{aa} \quad \text{and} \quad \mu_{bb} > \mu_{ab} > \mu_{aa}.$$

The joint distribution is of θ, given by $q = (q_{aa}, q_{ab}, q_{ba}, q_{bb})$, is depicted in Table 1. Here $q_1 =$

		θ_2: a	θ_2: b	
θ_1	a	q_{aa}	q_{ab}	q_1
	b	q_{ba}	q_{bb}	$1-q_1$
		q_2	$1-q_2$	

Table 1: Joint distribution of θ.

$q_{aa} + q_{ab}$ and $q_2 = q_{aa} + q_{ba}$ are the marginal distributions of θ_1 and θ_2, respectively. Let ρ be the correlation between θ_1 and θ_2, and define $\sigma = \sqrt{q_1(1-q_1)}\sqrt{q_2(1-q_2)}$. As shown in Table 2, the distribution can then be rewritten using three parameters: ρ, q_1, q_2.

The insuree observes θ_1 and knows the marginal distribution of θ_2, and the insurer knows the joint distribution of θ—this non-standard endowment of information in an insurance problem is what leads to the inverse selection structure. In terms of the primitives, we assume that q_1 and q_2 are common knowledge, the agent is privately informed about θ_1, and the principal privately knows ρ. Finally, to close the model, we assume that ρ is drawn from F on $[\underline{\rho}, \overline{\rho}]$, where F is differentiable, has a continuous density f, and is commonly known.[14]

[13] A standard behavioral foundation for the mean-variance preference is the CARA-Gaussian model, which has been used in many seminal papers, including Grossman and Stiglitz [1980].

[14] The entire set of possible correlation is of course [-1,1]. However, once we fix the marginals to be q_1 and q_2,

		θ_2		
		a	b	
θ_1	a	$q_1 q_2 + \rho\sigma$	$q_1(1-q_2) - \rho\sigma$	q_1
	b	$(1-q_1)q_2 - \rho\sigma$	$(1-q_1)(1-q_2) + \rho\sigma$	$1-q_1$
		q_2	$1-q_2$	

Table 2: Joint distribution of θ in terms of correlation.

The question we ask is: what is the principal optimal contract in this insurance problem?

Remarks on modeling. We intentionally model the distribution of information between the insurer (principal) and insuree (agent) as the former knowing ρ and latter knowing θ_1 to capture the idea that the insurer has some statistical knowledge and the insurer has some concrete knowledge about the underlying state. After the endowment of initial information, the insurer knows more about the general environment in the form of the correlation coefficient between the two dimensions, and the insuree knows something specific about his situation in the form of θ_1. Once the insurer incentivizes the insuree to reveal θ_1, the insurer can make better inference about the state than the insuree. This "inverts" the selection problem.

2.2 The optimization problem

To write down the problem formally, we introduce the associated mechanism design lexicon in the spirit of Myerson [1982, 1983]. A message rule $r : [\underline{\rho}, \overline{\rho}] \to \Delta(\mathcal{M})$ represents how coarsely (or finely) the insurer wants to communicate her information about the correlation coefficient to the insuree, as part of the optimal contract. For example, if only two contracts are offered, the set $[\underline{\rho}, \overline{\rho}]$ is partitioned into two sets, hence two messages are produced, one corresponding to each contract. So the agent knows, upon seeing one of the contracts which of the two sets the true ρ is placed in and updates her prior accordingly. If r instead assumes only one value, no information beyond the prior is communicated or if r assumes strictly monotone values, then information is fully disclosed.

Further, invoking the revelation principle, we simply look at a direct mechanism where the insurer reports her "type" ρ, the insuree reports his "type" θ_1, and a contract is selected from the menu:

$$C = \left\{c_m(a), c_m(b)\right\}_{m \in \mathcal{M}}, \text{ where } c_m(\theta_1) = \left(p_m(\theta_1), x_m(\theta_1)\right) \text{ for } \theta_1 = a, b.$$

A direct mechanism is then completely captured by (r, C), which is chosen by a *mediator* with the objective of maximizing the profit of the insurer subject to incentive compatibility and individual

it can be easily checked that the set of feasible correlations is $[\underline{\rho}, \overline{\rho}]$, where $\overline{\rho} = \min\left\{\frac{q_1(1-q_2)}{\sigma}, \frac{q_2(1-q_1)}{\sigma}\right\}$ and $\underline{\rho} = \max\left\{-\frac{q_1 q_2}{\sigma}, -\frac{(1-q_1)(1-q_2)}{\sigma}\right\}$. Thus, the support of F is restricted by the marginals q_1 and q_2.

rationality for the insuree, and potentially the incentive compatibility for the insurer.

The exact timing of the (dynamic) mechanism is as follows.

Stage 1	Stage 2
• nature draws $\rho \sim F \wedge \theta \sim \mathbf{q}$. • insurer learns ρ and reports it. • r generates message m. • insuree forms posterior F_m.	• menu $\{c_m(a), c_m(b)\}$ is offered. • insuree learns θ_1 and reports it. • contract $c_m(\theta_1)$ is implemented. • payoffs π and u are realized.

The goal going forward is to characterize the optimal choice of (r, C). To that end, we now define the objective and constraints of the optimization problem. A conceptual (and modeling) choice here emerges in how much commitment power the principal must have on the disclosure of information. We will start with the canonical information design problem layered on top of the mechanism design problem, wherein the principal chooses a messaging rule r at the outset, before ρ is realized. Beyond its reasonableness as a theoretical benchmark, a simple motivation for this is regulatory diktats and reputational concerns about not using private information to tailor contracts each time, independently.[15]

Let $\pi(\rho)$ be the (ex post) profit of the insurer if her type is ρ. Then the (ex ante) objective of the mechanism design exercise is given by:

$$\Pi = \int \pi(\rho) f(\rho) d\rho.$$

For a fixed menu c_m, the payoff of the insuree type $\theta_1 \in \{a, b\}$ from reporting $\hat{\theta}_1$ is:

$$\begin{aligned} u_m(\hat{\theta}_1; \theta_1) &= w - p_m(\hat{\theta}_1) - \left[1 - x_m(\hat{\theta}_1)\right] \mu_m(\theta_1) - \frac{\eta}{2}\left[1 - x_m(\hat{\theta}_1)\right]^2 \\ &= \underbrace{w - \mu_m(\theta_1)}_{\varphi_m(\theta_1)} + \underbrace{\left[x_m(\hat{\theta}_1)\mu_m(\theta_1) - \frac{\eta}{2}\left\{1 - x_m(\hat{\theta}_1)\right\}^2\right]}_{v_m(\theta_1; \hat{\theta}_1)} - p_m(\hat{\theta}_1) \\ &= \varphi_m(\theta_1) + v_m(\theta_1; \hat{\theta}_1) - p(\hat{\theta}_1) \end{aligned} \tag{1}$$

where $\mu_m(\theta_1)$ is the expected value of μ based on the realized value of θ_1 and the insuree's beliefs about ρ after observing the message m. The mathematical expression for the insurer's profit is:

$$\pi(\rho) = q_1 \left[p_{r(\rho)}(a) - \mu_\rho(a) x_{r(\rho)}(a)\right] + (1 - q_1)\left[p_{r(\rho)}(b) - \mu_\rho(b) x_{r(\rho)}(b)\right] \tag{2}$$

where $\mu_\rho(\theta_1)$ is the expected value of μ based on realized value of ρ and (truthfully) reported value of θ_1, and $r(\rho)$ is the message m plugged in equation (1) above.

[15] See Best and Quigley [2023], Fréchette, Lizzeri, and Perego [2022], and Lin and Liu [2023], amongst others, for different foundations of the commitment assumption in persuasion and information design problems.

Two types of constraints are imposed on the optimization problem. First is the incentive constraint for the insuree (agent), that the insuree wants to truthfully report his type to the mediator:

$$IC_{\theta_1} : u_m(\theta_1; \theta_1) \geq u_m(\theta_1; \hat{\theta}_1) \; \forall \; \hat{\theta}_1.$$

As pointed out in the description of the dynamic mechanism above, insuree's incentive constraint incorporates the message coming from (technically the mediator through the mapping r), by conditioning the (expected) utility on m, and hence the posterior F_m. Second is the individual rationality constraint of the insuree which guarantees him a minimum expected utility:

$$IR_{\theta_1} : u_m(\theta_1; \theta_1) \geq u_0,$$

where u_0 is calculated by substituting $x = p = 0$ in Equation (1). The information plus mechanism design problem then reads as follows:

$$(\mathcal{P}) : \quad \max_{r,C} \Pi \text{ s.t. } IC_{\theta_1}, IR_{\theta_1}.$$

3 Benchmark: ρ is common knowledge

3.1 Rothschild-Stiglitz benchmark with a twist

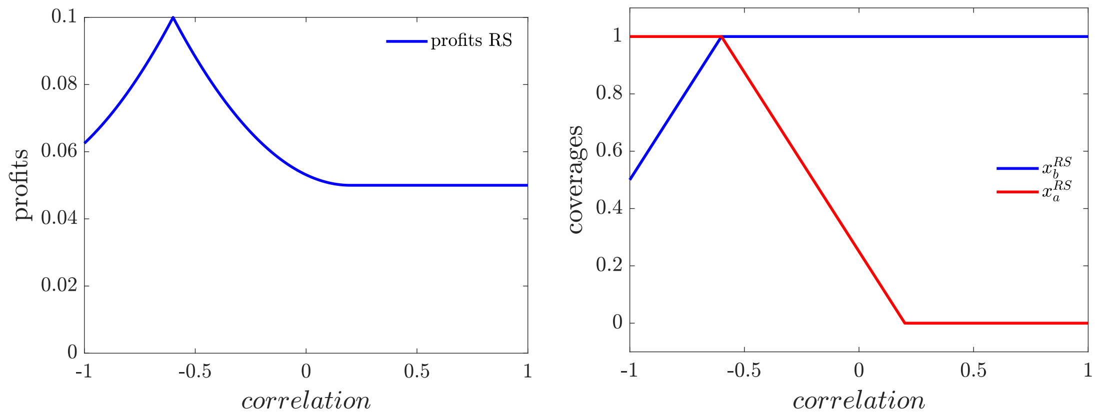
We start the analysis by looking at the benchmark model where ρ is common knowledge. So the principal has no statistical informational advantage, and hence there is no scope for inverese selection. The problem becomes isomorphic to the monopolistic version of the classical Rothschild and Stiglitz [1976] problem, studied first by Stiglitz [1977]. But, there is also a twist—it depends on ρ whether $\theta_1 = b$ or $\theta_1 = a$ is the "high" risk type. Let H denote the high risk type, and L the low risk type.

When ρ is common knowledge, there is no need of communication from the insurer; hence there is no role for $r : [\underline{\rho}, \overline{\rho}] \to \Delta\mathcal{M}$. Both parties take expectations over θ_2 using the prior, and insuree is incentivized to reveal θ_1 truthfully. We will refer to this as the *benchmark model*, and label it RS, pointing to the classical reference. The optimal contract is as follows:

Proposition 1. *The optimal coverage is* $1 = x_\rho^{RS}(H) \geqslant x_\rho^{RS}(L) \geqslant 0$, *where* $\exists\, \rho^* \in [\underline{\rho}, \overline{\rho}]$ *that solves* $\pi^{RS}(\rho^*) = \max_\rho \pi^{RS}$ *s.t.*

1. $\rho > \rho^* \Rightarrow H = b, L = a,$
2. $\rho < \rho^* \Rightarrow H = a, L = b.$ *and*
3. $\rho = \rho^* \Rightarrow x_\rho^{RS}(b) = x_\rho^{RS}(a)$

As in the standard monopolistic screening model, the optimal contract is always separating: "high" risk type is offered exact coverage and "low" risk type is offered partial coverage, though which type is high risk pivots around ρ^*. Fix ρ^* to be the correlation where the expected value of



(a) profit as a function of ρ (b) coverage as a function of ρ

Figure 1: Benchmark model when ρ is common knowledge

mean loss is the same for both θ_1-types: that is, ρ^* solves $\mu_\rho(b) = \mu_\rho(a)$. Then, for $\rho > \rho^*$, the high risk type is $\theta_1 = b$ and for $\rho < \rho^*$, the high risk type is $\theta_1 = a$ (see Figure 3b). The profit is maximized at ρ^*, because the agent's private information of θ_1 becomes statistically irrelevant: the principal offers a pooling contract and extracts all the surplus associated with it (see Figure 3a).[16] Naturally, the profit function is also single peaked at ρ^* for in either direction, the private information of the insurer has becomes more precise as ρ gets further away from ρ^*. The two pieces are convex, i.e. the drop in profit happens at a decreasing rate—initially the rate of reduction is high and it gets progressively slower.

For the case $\rho > \rho^*$, for an interior solution, the optimal coverages are given by:

$$x_\rho(b) = 1 \text{ and } x_\rho(a) = 1 - \frac{1-q_1}{\eta q_1}(\mu_\rho(b) - \mu_\rho(a)) < 1 \tag{3}$$

As can be seen transparently, the extent of distortion for the "low" risk type is further determined by the primitives of the problem. In particular, the distortion is decreasing in η and q_1, and increasing in $\mu_\rho(b) - \mu_\rho(a)$. Also for a large enough distortion, the "low" type is shut out of the market, as seen in Figure 3b for high values of ρ. Analogous comparative statics emerge for the case $\rho < \rho^*$.

The novel twist here is that the notion of a high or low risk type is endogenous to realization of ρ. This already indicates what may happen when ρ is privately observed by the insurer: She will have two motivations then. First, is to make the agent believe that the realized type is ρ^* or at least in its neighborhood. And the second force is that the insurer would like to mislead the agent: when ρ is large and hence $\theta_1 = b$ is the high type, the insurer would like the insuree to believe the realization of ρ is small, so he wrongly deduces that $\theta_1 = a$ is the high type, and so on. Both these forces will in turn be tempered by the insuree's ability to perform Bayesian inference on observing and interpreting the set of offered contracts.

[16]Technically speaking, for $\rho > \rho^*$, IC_b binds at the optimum, and for $\rho < \rho^*$, IC_a binds at the optimum. This determines which type is offered the efficient contract and which one is distorted.

3.2 What happens when ρ is private information of the principal?

In all the models that follow, ρ is not common knowledge, rather it is the insurer's private information. This will feature an inversion of adverse selection: by designing an incentive compatible mechanism, once the insurer learns θ_1, she knows more than the agent about the probability of the state θ.

The distinct question for contract design here is how finely should the offered contract depend on the value of ρ? The insurer faces a new trade-off: On the one hand, the more finely the contract depends on ρ, the greater *price discrimination* can be achieved. On the other hand, the more finely the contract depends on ρ, the more information the insuree (agent) can extract about ρ. Bayesian inference allows her to update the prior about principal's ρ from seeing the offered contract. The insurer wants to maintain a belief gap by *obfuscating* the exact value of the correlation ρ and hence the true probability of loss. We call this trade-off the *price discrimination* versus *obfuscation* trade-off. It exists in addition to the traditional *rent* versus *efficiency* trade-off described above.

In what follows, we first present the optimal information and contract design problem, followed by relaxing the insurer's decision to choose the information structure at the outset. Following this, we describe the necessity of exploring behavioral variations of our model, where the insuree cannot do Bayesian inference properly. To that end, we present two such models, and discuss the contracts that then emerge.

4 Optimal information and contract design

In this section we introduce statistical information advantage for the insurer. She learns ρ, which captures the new asymmetry of information made possible due to modern techniques of big data and machine learning. Once the insurer elicits the information of the insuree, inverse selection can emerge, wherein the insurer can exploit her informational advantage.

4.1 Full commitment solution

A first step towards understanding the behavior of a rational insuree and associatedly a profit maximizing insurer is study the joint and information and contract design problem marked ($\mathcal{P}$) above. Here the insurer chooses the information disclosure policy at the outset, a a function of the correlation coefficient, and then offers a screening contract conditional on the information disclosed. The insuree updates his beliefs according to Bayes' rule and incentive and individual rationality constraints are evaluated using this posterior.

Following the burgeoned literature on Bayesian persuasion and information design (Kamenica and Gentzkow [2011] and Bergemann and Morris [2016]), the principal or the insurer's problem can be considered as choosing posterior beliefs that satisfy the martingale (or Bayes' plausibility) condition, and then which contract to offer at each of those posterior beliefs. Since both the insurer and insuree's payoffs are linear in ρ, the object of information design, it is sufficient to consider only the expectation generated by the posterior distribution of the disclosure policy (Gentzkow and

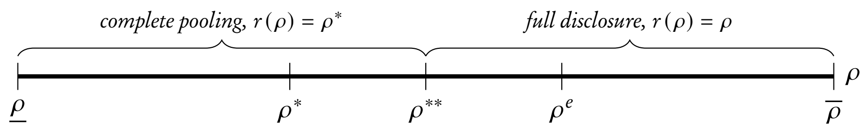
Kamenica [2016]). Thus, without loss of generality, we can recast the problem so that $\mathcal{M} = [\underline{\rho}, \overline{\rho}]$, and $r : [\underline{\rho}, \overline{\rho}] \to \Delta\mathcal{M}$ maps each realization of ρ to its targeted expected correlation.

Once the message is sent, the insuree (or the agent) forms a posterior with mean say $\hat{\rho}$. It is then easy to see that the contract design problem is akin to the benchmark RS problem we studied in the previous section at the correlation $\hat{\rho}$ where the insurer's optimal profit is given by $\pi^{RS}(\hat{\rho})$. As in Figure 3a, this function is convex at both sides of ρ^* and has a global maximum at ρ^*.

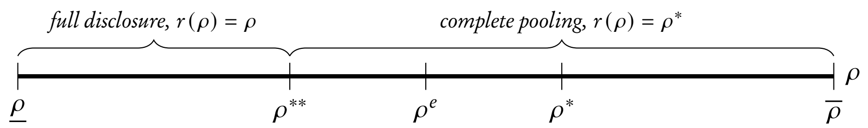
So the solution to problem $(\mathcal{P})$ boils down to what set of posterior beliefs that satisfy Bayes' rule will maximize the seller's expected profits given that the RS-contract will be implemented at each of those (expected) beliefs in the second, contract design, stage. The solution of this problem is conceptually simple. The following proposition summarizes it.

Proposition 2. *Suppose $\rho^* \in (\underline{\rho}, \bar{\rho})$ and let $\rho^e = \mathbb{E}[\rho]$ be the ex-ante mean. The optimal information structure designed by the insurer is as follows:*

- *If $\rho^e = \rho^*$, pool all ρ-types by sending a unique message, $r(\rho) = \rho^*$.*
- *If $\rho^e > \rho^*$, $\exists\ \rho^{**} > \rho^*$ such that for all $\rho < \rho^{**}$, pool all ρ-types by sending one message $r(\rho) = \rho^*$, and for $\rho > \rho^{**}$, disclose all information by sending messages $r(\rho) = \rho$.*

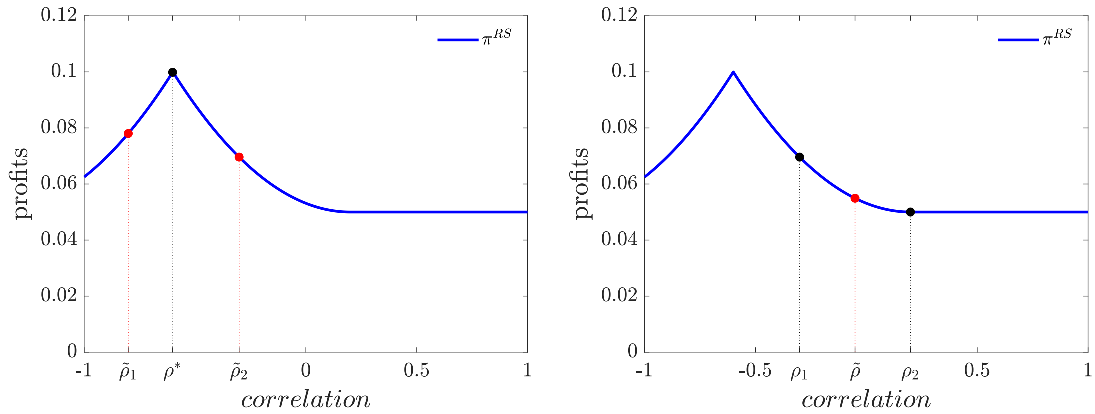


- *If $\rho^e < \rho^*$, $\exists\ \rho^{**} < \rho^*$ such that for all $\rho > \rho^{**}$, pool all ρ-types by sending one message $r(\rho) = \rho^*$, and for $\rho < \rho^{**}$, disclose all information by sending messages $r(\rho) = \rho$.*



In plainer words, the insurer wants to push as much mass as possible to the point ρ^*, to the extent permitted by the Bayes' plausibility condition. This is because the insuree has the least informational advantage, and thence the insurer has her maximum profit at ρ^*. This culminates in pooling of types around ρ^* in a way that the posterior mean is exactly ρ^*. If the ex ante mean is also ρ^*, that is the best possible scenario for the insurer, so r is simply a constant map that doesn't disclose any information beyond the prior. In all other cases, when $\rho^e \neq \rho^*$, the pooling region is determined by whether the ex ante mean lies to the left or the right of ρ^*. In the non-pooling region, forced onto the insurer to satisfy the Bayes' plausibility condition, it is optimal to do full information disclosure due to the convexity of the profit function.

Figure 2 provides some graphical intuition. Suppose only two types, ρ_1 and ρ_2, are possible. In Figure 2a, they lie on different sides of ρ^*. It is visually easy to see that the expected profit of offering the RS contract at each of these values individually is dominated by the RS profit at the



(a) Posterior beliefs on both sides of ρ^* (b) Mixing two correlations at one side of ρ^*

Figure 2: Intuition for Proposition 2

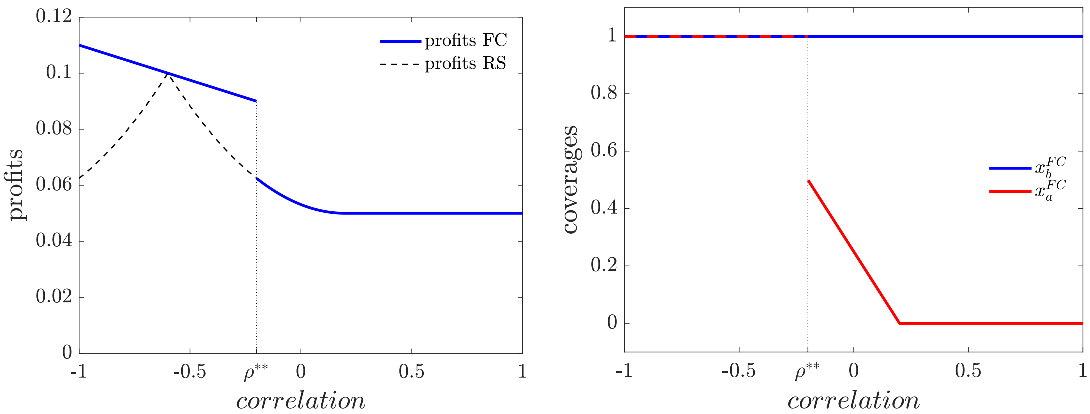
expected value of ρ from the two points ρ_1 and ρ_2. So, the insurer will always pool these types. In fact the insurer wants to do more and try to choose signals so as to get the mean as close to ρ^* as possible. Broadly, the same intuition holds if $\hat{\rho}_1$ and $\hat{\rho}_2$ are posterior means instead of two discrete types.

Next, consider the case when ρ_1 and ρ_2 are the only feasible correlations and on the same side of ρ^*, as in Figure 2b. Since π is strictly convex in this region, a simple application of Jensen's inequality will tell you that the expected profit on offering a separate contract at ρ_1 and ρ_2 will be higher than a single contract at $\hat{\rho}$, the expected value of ρ_1 and ρ_2. Hence, the insurer always offers a separating contract. Again, taking this intuition to the general continuous type space: strict convexity of the profit function ensures that full separation improves upon any pooling of types on one side of ρ^*.

Putting these facts together, the insurer pools the types in the neighborhood of ρ^*; in fact, pools all types of one side of the peak, and exact location of ρ^e with respect to ρ^* determines which side is pooled, and types on the other side of ρ^* are also pooled as much as possible, to extent permitted by the Bayes' plausibility condition. The posterior mean of this pooled mass must exactly equal ρ^*. What remains of the type space has full disclosure, due to the convexity of the profit function. Figure 3a depicts the profit function, and how a discontinuity thus arises because of the from taken by the optimal information structure.

As a summary, the inverse selection problem has a "bang-bang" sort of solution: in one region where the principal is expected to gain large profits by hiding the true information, she maximally obfuscates while respecting the ability of the insuree to do Bayesian inference correctly, and in the other region, she discloses it all to do maximal price discrimination.

Finally, for the contract design part, as already discussed, at each realized posterior ρ, the insurer simply offers the RS contract at the correlation. Figure 3 depicts the pointwise profit and coverage in the optimal contract as a function of ρ. In the pooling region, for $\rho < \rho^{**}$, the RS contract corresponding to correlation ρ^* is offered, which is full coverage. Since the insuree's expected correlation is ρ^*, he has no statistical advantage whatsoever, and hence full surplus is extracted.



(a) profit as a function of ρ (b) coverage as a function of ρ

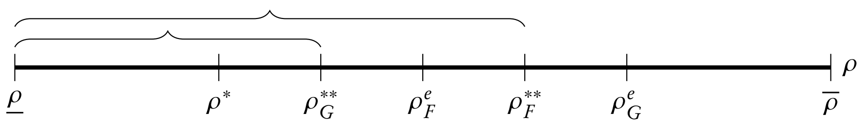
Figure 3: Optimal contract and profits with full commitment

And π is a linear and decreasing function of ρ, so in the pooling region, π is decreasing and always above the π^{RS}. In the full disclosure region, both the profit and coverages are exactly the same as the RS benchmark.

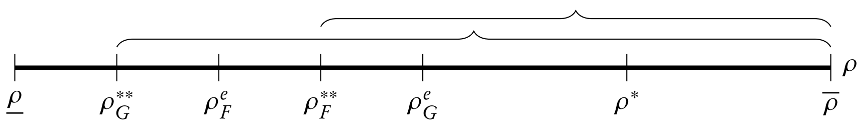
A natural comparative statics question to ask is how does the set of correlations for which we see complete pooling and full disclosure change as the underlying prior distribution changes. To that end, we invoke the notion of conditional stochastic dominance: A distribution function G conditionally stochastically dominates F if it has less probability mass in every right truncated left tail (see Riley [2017]). Note that when G conditionally stochastically dominates F, then $\rho_G^e = \mathbb{E}_G[\rho] \geqslant \mathbb{E}_F[\rho] = \rho_F^e$.

Proposition 3. *Suppose that the continuous distribution G conditionally stochastically dominates the distribution F, that is, $\frac{G\rho)}{G(\tilde{\rho})} \leq \frac{F(\rho)}{F(\tilde{\rho})}$ for all $\tilde{\rho} \in [\underline{\rho}, \bar{\rho}]$ and $\rho \leq \tilde{\rho}$.*

1. *If $\rho_F^e > \rho^*$, the set of correlations with complete pooling for F is larger than that for G.*



2. *If $\rho_G^e < \rho^*$, the set of correlations with complete pooling for G is larger than that for F.*



Conceptually, the comparative static shows whether more or less set of types are included in the pooling or full disclosure region. Consider the first case above when both expectations, ρ_F^e and ρ_G^e, lie to the right of ρ^*. So from Proposition 2 we know that the full disclosure region lies on the right side of the interval and pooling towards the left, split in each case by the corresponding ρ^{**}. Since G conditionally stochastically dominates F, the conditional expectation of the right truncated distribution at any correlation ρ is smaller for F than for G. As the conditional expectation of the correlations that the insurer pools must be equal to ρ^*, a larger length of the support of correlations needs to be included in F than is G, which delivers the comparative statics result. Analogous arguments hold for the second case.[17]

4.2 Limited commitment model

We made one important assumption in Section 4.1 that runs through the information design literature: the sender or the principal (in this case the insurer) chooses the information structure or disclosure policy r at the outset, before ρ is realized. This is, we think, both a reasonable theoretical benchmark and arguably realistic when regulatory or reputational concerns prevent her from tampering with how to tailor the information in each contract.

However, it also true that in some specific circumstances, the insurer cannot be stopped from switching to a different strategy of how to use the information once she has learnt it. To model this limited commitment in the information design problem, we use the standard approach of imposing an incentive constraint for the principal:

$$IC_\rho : \pi(\rho;\rho) \geq \pi(\rho;\hat{\rho}) \;\forall\; \hat{\rho},$$

where the profit function, $\pi(\rho;\hat{\rho})$ for realized type ρ and reported type $\hat{\rho}$, is given by

$$\pi(\rho;\hat{\rho}) = q_1\left[p_{r(\hat{\rho})}(a) - \mu_\rho(a)x_{r(\hat{\rho})}(a)\right] + (1-q_1)\left[p_{r(\hat{\rho})}(b) - \mu_\rho(b)x_{r(\hat{\rho})}(b)\right]. \tag{4}$$

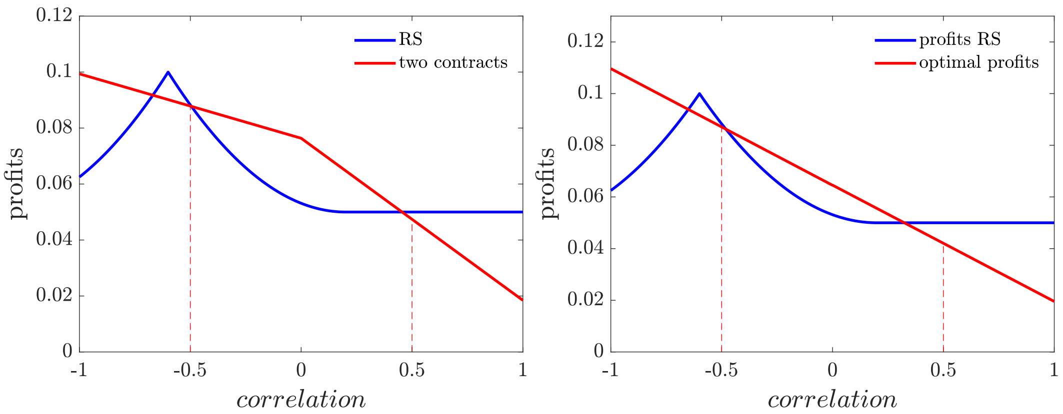
So the limited commitment information and contract design problem reads as:

$$(\mathcal{P}-LC): \quad \max_{r,C} \Pi \text{ s.t. } IC_\rho, IC_{\theta_1}, IR_{\theta_1}.$$

Conceptually, this makes the information design problem more constrained, so that the obfuscation versus price discrimination trade-off must shift more in favor of the obfuscation because presumably less information will be revealed than in the commitment model. Technically, we show this is indeed the case, though this is a significantly harder problem for as in the classic cheap talk model (Crawford and Sobel [1982]), we have to simultaneously determine the number of partitions of the type space and the placement of these partitions. In addition, there is a contract design problem associated with each element of such a partition.

Recollect from standard Myerson mechanism design (Myerson [1981]) that an incentive constraint is equivalent to an envelope theorem and a convexity (or monotonicity) constraint, and whenever the convexity constraint is violated by the allocation satisfying only the envelope condition, the allocation rule is *ironed*. In the appendix we show that the incentive constraint on the side of the insurer here, IC_ρ, is such that the allocation rule, that is the coverage rate x, is ironed

[17] Note that when $\rho^e_F < \rho^* < \rho^e_G$, a global relationship between the pooling (or full disclosure) regions under F and G is not possible, and it will depend on the fine details of the parameters.



(a) optimal profit ignoring convexity (b) optimal two element partition

Figure 4: Optimal partitions

almost everywhere on the type space.

Ironing almost everywhere immediately implies that the optimal information structure, captured by the mapping r, is partitional, and only a finite number of messages are sent, and hence a finite number of posterior beliefs are formed by the insuree. In fact, in most parametric settings, only one or two messages are sent, hence our claim that the trade-off between belief gap and price discrimination is now firmly resolved in favor of the belief gap.

The basic intuition from the previous section carries through to the extent that the insurer wants to push as much mass around ρ^* towards it, and pool all those types. But the limited commitment assumption brings another key force: the insurer upon learning the actual correlation would like to misreport it, and sell a lot of insurance to the low risk insuree and little to the high risk insuree. Being rational, the insuree can also deduce this temptation of the insurer and work backwards from it. This newer force therefore overturns the full disclosure part of Proposition 2, culminating in very few contracts being offered at the optimum.

The main theoretical result from this section are now stated, the formal proofs can be found in the appendix.

Proposition 4. *Suppose the design problem also needs to satisfy IC_ρ. Then the following hold:*

- *a finite number of contracts are offered at the optimum, hence $|\mathcal{M}| = |C| = \kappa < \infty$.*
- *as the underlying risk gets arbitrarily large, only one contract is offered at the optimum.*

Solved examples and numerical results further suggest that the optimal number of contracts is not just finite but at most two, that is $|\mathcal{M}| = |C| \leqslant 2$. A brief graphical intuition for ironing and hence the fewness of contracts is offered in Figure 4. The blue single peaked curve represents the RS profits when ρ is common knowledge. The red line depicts the optimal profit curve. For the chosen parameters, the optimal information structure produces a two-element partition of the type space.

The expected correlation for each partition in Figure 4 is marked by the dotted vertical lines. Feasibility demands that the red profit line must not be above the blue benchmark profit at each of those two points. This is because in the subgame in which correlation is common knowledge, the best the insurer can do is to achieve a profit of $\pi^{RS}(\rho)$. In the two subgames, one for each of the two partitions, it is as if the insurer is in the benchmark model with the correlation being the expectation of correlation in those partitions.

Recollect that the profit function is linear in ρ. So, respecting feasibility, we could choose the highest piecewise linear curve that crosses the blue curve at those expected correlations. Solving for the optimal contract ignoring the convexity constraint would culminate in a concave kink in the piecewise linear profit function, as shown in Figure 4a. Thus, the ignored convexity constraint is violated, and this observation is not limited to the parameters chosen here—*any* two (or more) partition contract which solves the relaxed problem will generate such a concave kink. Thus, to make the contract incentive compatible, it must be *ironed*. The highest convex profit function that the insurer can construct while satisfying incentive-feasibility is the straight red line in Figure 4b.

Unlike the full commitment model, the profit is not uniformly higher than the RS contract, compare Figure 4b with Figure 3a. In fact, with limited commitment, ironing forces the profit curve to be below the RS profit at the expected correlation of the second (or right) partition of the optimal information structure. This represents the cost of the incentive constraint for the principal's profit and also for the total surplus.

Finally, if we increase the number of partitions from two to three, we would draw the best piecewise linear function that is convex and weakly below the benchmark profit at each of the three expected correlations corresponding to the three partitions. The costs of doing this outweigh the benefits, and this intuition holds more generally: At the optimum, the principal does not want to have numerous contracts because the costs of distortions introduced to satisfy incentive compatibility across partitions outweigh the benefits accrued from greater price discrimination. Thus, the number of contracts offered is not just finite, it is quite small. The fact that the insuree can do perfect Bayesian inference washes away most of the advantage of the insurer to use the information on ρ.[18]

5 Behavioral approach: incomplete inference from contracts

5.1 Basic motivation

In the last section, we solved for the optimal information structure and the associated screening contract when the agent, that is the inusree, can do perfect Bayesian inference, and the principal, i.e. the insurer, can pick the information disclosure policy and contract space at the outset. This produces a stark result that the insurer does one of two things: In a substantial part of the type space, she forgoes her informational advantage by providing a unique contract pair of premium and coverage, and in the rest of the type space, she does full disclosure, and tailors the contract precisely to her private information.

[18] In the interest of expositional simplicity, we do not discuss the exact structure of the contract here, this is presented in the appendix.

So, the inverse selection problem leads to one of two extreme outcomes, in one the obfuscation versus price discrimination paradigm is resolved in favor of maintaining obfuscation and giving up on price discrimination, and in the second it is resolved in favor of price discrimination by giving up on obfuscation completely. Relaxing commitment to the information structure by the insurer leads to an even more stringent outcome, in that the trade-off is resolved more firmly in maintaining obfuscation and not using the information on ρ to price discriminate too much.

While these are informative benchmarks, the requirement of Bayesian inference here by the insuree is strong. There is plenty of evidence in field and laboratory settings that show humans find it hard to fully do Bayesian updating when faced with actual economic problems; see Benjamin [2019] for extensive overview of the literature. In the context of our model the insuree's ability to do Bayesian inference works at two levels. First, it is about understanding or interpreting how to map the space of offered contracts to information revelation and second,the Bayesian updating of probabilities itself. This extra first step has its own considerable complexity in a real marketplace. In this spirit, Handel and Kolstad [2015] argue that standard insurance models and empirical analysis of health insurance markets based on these models miss this difficulty: "Traditional models of insurance choice are predicated on fully informed and rational consumers protecting themselves from exposure to financial risk. In practice, choosing an insurance plan is a complicated decision often made without full information."

In addition, the rise in data markets and data intermediaries in the real world, tells us that the returns to inverse selection for the principal, i.e., a firm with access to big data technology, are potentially large. In the context of our model, however, Bayesian inference washes away, at least in the limited commitment version, a substantial chunk of the informational advantage that the insurer can exploit, as is the case in standard informed principal problems.[19] Presumably, in settings with limited inference capacity of the agent, the price discrimination channel can be exploited more without compromising too much on obfuscation.

As a motivation to address these issues, we explore two simple yet extreme behavioral models in which the insuree cannot do Bayes inference correctly. In the first, the insuree is gullible to the extent that he believes the correlation ρ told to him by the insurer. And, second, the insuree simply sticks to the prior and does not update no matter what contract is offered by the insurer. In both cases, we study the contract design under those specific information paradigms. At the end of the section, we discuss the key forces driving these results. Note that in both discussed cases the solution to the commitment and limited commitment problem is the same, so we don't emphasize the distinction.

5.2 Gullible insuree

A rather non-standard model to consider is one where, in addition to offering a contract, the insurer tells the insurer the correlation coefficient and the insuree simply believes it. This setting

[19] In fact Myerson [1983] calls the *inscrutability principle* the idea that the principal will not use her information while offering contracts because it will be detected in equilibrium by the agent.

is different than the (standard) naïveté model that we discuss in the next subsection. We will refer to such an insuree as gullible. This allows the insurer maximal freedom in shaping the insuree's beliefs and tailor the contract then to those beliefs. We don't necessarily think of this as a realistic model, but is instructive in showing the direction in which the information and contract design will go if the insurer has sharp instruments at her disposal to shape the insuree's beliefs that do not have to satisfy the Bayes' plausibility condition.

Knowing that she can basically mislead the insuree about the way in which the two dimensions are correlated provides the insurer with great freedom in selecting contracts. She will choose r and C in tandem to create both the maximal obfuscation and the maximal price discrimination.[20]

Proposition 5. *If the insuree is a gullible, $\exists\ \hat{\rho} \in [\underline{\rho}, \overline{\rho}]$ such that:*

1. *binary messages sent: $\mathcal{M} = \{\underline{m}, \overline{m}\}$ s.t. $r(\rho) = \overline{m}$ for $\rho < \hat{\rho}$ and $m(\rho) = \underline{m}$ for $\rho > \hat{\rho}$,*
2. *insuree's posterior is extreme: $F_{\underline{m}} = \delta_{\underline{\rho}}$ and $F_{\overline{m}} = \delta_{\overline{\rho}}$ where δ_ρ is Dirac measure on ρ,*
3. *profits are uniformly higher than benchmark: $\pi(\rho) > \pi^{RS}(\rho)\ \forall\ \rho$ almost surely,*
4. *coverages are generically separating and inefficient: $x_\rho(H) \neq x_\rho(L)$ and $x_\rho \neq 1\ \forall\ \rho$ a.s.*

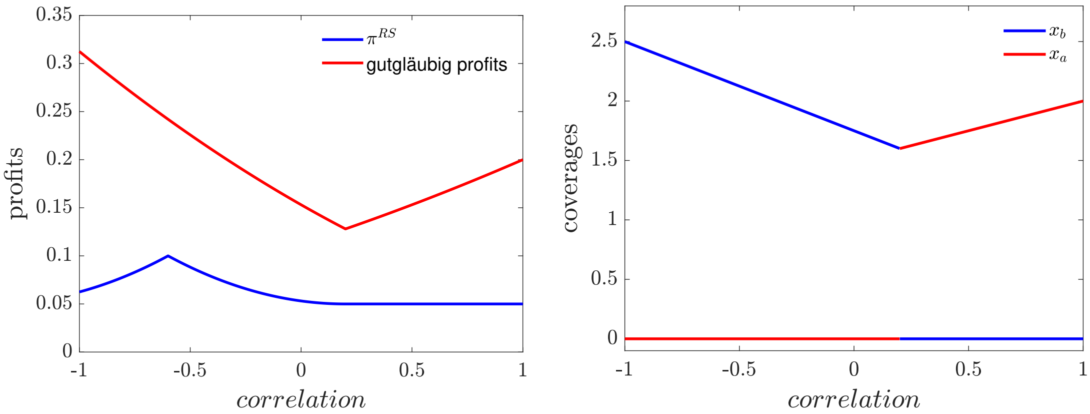
There exists a threshold value of ρ, to the right of which the insuree reports the extreme negative correlation, $\underline{\rho}$, and to the left of which she reports the extreme positive correlation, $\overline{\rho}$. Even though the cardinality of the message space is just 2, a distinct contract is offered for each value of ρ, since the insuree does not infer anything about ρ from the menu of contracts.

When the *actual* correlation is high, it means that the type $\theta_1 = b$ is likely to suffer a large loss and $\theta_1 = a$ is likely to suffer a small loss. In this scenario, the insurer *reports* a large negative correlation, in fact the largest possible negative value, and overinsures $\theta_1 = a$ and underinsures $\theta_1 = b$. In the process, she is able to achieve dramatic price discrimination while ensuring maximal obfuscation. The exact opposite is true for the case when the actual correlation is low: insuree reports largest positive correlation, and overinsures $\theta_1 = b$ and underinsures $\theta_1 = a$. In sum, the insurer sells a large amount of insurance at a high price to the type who actually has a low probability of loss, and a small amount of insurance to the type who actually has a high probability of loss.

Figure 5 depicts the profit and coverages when the insuree is gullible. That the profits are uniformly higher (Figure 5a) is intuitive—the model allows the insurer to input any value of the correlation in the insuree's incentive compatibility condition, which in turn allows her to manipulate which type is perceived to be high risk type and then decide what coverages to offer each θ_1 type (Figure 5b).[21]

[20] Since the Bayes' consistency condition is not valid, technically the class of contract is given by $C = (c_{m,\rho})$ because the contract offered for the actual realization of ρ has no relation to the reported value m.

[21] The overinsurance offered at the extremes brings out the message starkly. One can limit the coverage exogneosuly to be below one, that is, $x \leq 1$. Whenever it is optimal to set $x > 1$, the bound will be hit and will get full but not over insurance. All other results will continue to hold qualitatively.



(a) profit as a function of ρ (b) coverage as a function of ρ

Figure 5: Model with gutgläubig insuree

5.3 Naive insuree

A more standard "behavioral" way of modeling limitations on information processing is to assume that the agent ignores the signals offered by the contract about the correlation coefficient, so that $F_m = F \ \forall m \in \mathcal{M}$. Thus, in this situation, the role of r is moot. The insurer designs the contract as a function of ρ with the knowledge that the insuree will evaluate his payoffs using the prior F.

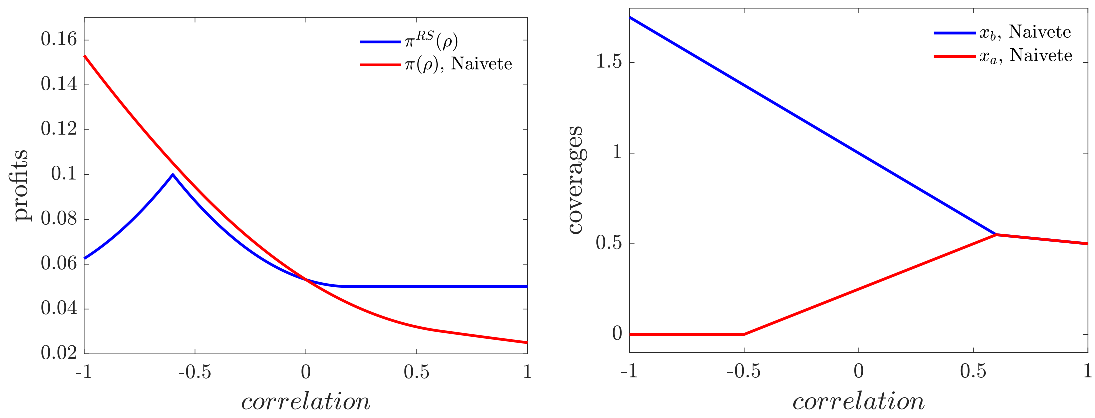
Proposition 6. *If the insuree is naive (and thus sticks to the prior):*

1. *profits are higher in expectation:* $\mathbb{E}(\pi(\rho)) > \mathbb{E}(\pi^{RS}(\rho))$,
2. *coverages features both pooling and separation,*
3. *coverages are generically inefficient:* $x_\rho(\theta_1) \neq 1 \ \forall \ \rho$ *a.s.,*

The salient difference between the naive model and gullible case (and also the general model discussed in Section 4) is that here the extent of obfuscation is determined exogenously by the fixed prior and the realization of ρ, and the insurer cannot influence it. This works in the insurer's favor sometimes, and other times it works against her. As a consequence, when the insuree is naive, the insurer is better off on average in comparison to the benchmark RS model, however, unlike the gullible case, this ranking is not uniform (see Figure 6a).

Here is a simple intuition for the result: Suppose the expected correlation according to F is high enough, so that the insuree thinks "high" risk type is $\theta_1 = b$. If the realized correlation is close to $\underline{\rho}$, the insurer wants to sell a lot of insurance to $\theta_1 = b$ and little insurance to $\theta_1 = a$. This is because $\theta_1 = b$ is actually the "low" risk type but he believes his risk to be at a higher level, according to F, and is happy to buy a lot of insurance (see left part of Figure 6b). On the other hand, when the realized correlation is close to $\overline{\rho}$, the insurer would like to sell a lot of insurance to $\theta_1 = a$, but the extent of it is limited by the agent's incentive constraint that demands $x_\rho(b) \geqslant x_\rho(a)$. Moreover, she cannot sell a lot of insurance to $\theta_1 = b$ either because the insuree does not internalize the extent of risk he faces. Thus, for extremely high correlations, the insurer is forced to pool the coverage.[22]

[22] If expected correlation according to F is low enough, then in a symmetric contrast to Figure 6, the profit curve



(a) profit as a function of ρ (b) coverage as a function of ρ

Figure 6: Model with naive insuree

5.4 Dissecting the key forces

A quick summary from all the cases discussed thus far:. The coverages vary as a function of ρ, the insurer's private information, and θ_1, the insuree's private information. The latter due to is the classical rent-versus-efficiency tradeoff which runs through each of the cases, since the insuree's incentive constrain needs to be satisfied. The former generates a distinct new tension between obfuscation and price discrimination.

In the benchmark Rothschild-Stiglitz case, when correlation is common knowledge, the extent of obfuscation is zero, and price discrimination is determined exogenously through the realized value of ρ. In the gullible case, both obfuscation and price discrimination are endogenously determined. Since the insurer can choose the contract independently of the insuree's belief, there is a no-longer a trade-off between obfuscation and price discrimination, and both are selected to maximize the insurer's profit. In the naïve case obfuscation exists but is determined exogenously as the insuree sticks to the prior no matter what contract is offered. Price discrimination is endogenously chosen to maximize the insurer's profit given the exogenous obfuscation constraint. This hurts the insurer in some part of the type space, but is profitable on average.

In the rational model, where the insuree can do perfect Bayesian inference, the trade-off between obfuscation and price discrimination is more stringent, and some part of the advantage from inverse selection for the insurer is washed away by the power of Bayesian inference. Even still, when the insurer can commit to information design, she can extract higher profits than the RS benchmark by pooling a lot of the information around ρ^* and maintaining obfuscation in some part of the type space, while disclosing full information everywhere else. However, when the insurer cannot commit to the information structure, very few contracts are offered since the added incentive constraint ensures the cost of segmenting the market and giving up on belief gap are too high in comparison to the benefits from price discrimination.

would intersect benchmark profits from below, and pooling in coverages will happen for high negative correlations.

6 Extensions

In this section, we explore two natural interventions to the information design model (studied in Section 4.1), which can be used to regulate the market with the aim of improving the 'consumer surplus'.[23]

6.1 Competition

Suppose there is an insurer who is informed about ρ, call it the big firm, and there is another insurer in the market that does not know this value and works with the prior F. Borrowing from the benchmark (in Section 3), the latter insurer is assumed to offer the Rothschild-Stiglitz contract evaluated at the expected value of ρ. The idea here is to introduce competition in a tractable way from regular firms who do not have the in-house expertise of big data.

To entice the insuree, the contract offered for each message sent by the big firm must provide the insuree a higher expected utility than the RS-contract offered by the other firm. Technically speaking this a strong IR requirement, and the rest of the problem remains the same as in Section 4.1.[24] The problem reads as follows:

$$\max_{r,C} \Pi \text{ s.t. } IC_{\theta_1}, IR^e_{\theta_1}$$

where r, C and IC_{θ_1} are as defined before, and $IR^e_{\theta_1}$ is modified so that the insuree's outside option is evaluated at the contract $c^e = \{p^{RS}_e(b), x^{RS}_e(b), p^{RS}_e(a), x^{RS}_e(a)\}$. Here $x^{RS}_e(b)$ and $x^{RS}_e(a)$ are defined in Proposition 1 at $\rho = \rho^e = \mathbb{E}[\rho]$, and $p^{RS}_e(b)$ and $p^{RS}_e(a)$ are the corresponding prices that maximize the insurer's expected utility. So, $IR^e_{\theta_1}$ reads as follows:

$$IR^e_{\theta_1} : u_m(\theta_1; \theta_1) \geq u^{RS}_e(\theta_1) \text{ for } \theta = a, b.$$

where $u_m(\theta_1; \theta_1)$ is defined in Equation (1) for $\hat{\theta}_1 = \theta_1$ and $u^{RS}_e(\theta_1)$ is the rent of type θ_1 at the expected correlation in the RS model.

At a first pass, Figure 7 presents the mechanical calculation of profits of the insurer in the benchmark RS problem where ρ is common knowledge, but for the lower red curve, the IR constraint takes the stronger form described in this section. So following Section 4.1, the information design problem now is performed the on this lower red curve as the full disclosure benchmark, instead of the upper blue curve. The same broad principles apply, and under some reasonable sufficient conditions a direct comparison can be made of the "producer", "consumer" and "total" surplus, with and without competition.

The formal statements and proofs are reported in the appendix. The high level message of the result is that introducing competition reduces total surplus and and insurer's profit but increases the payoff of the insuree. In fact, the decrease in total surplus is a direct consequence of the reduction

[23] Results for the introduction of the two interventions in the limited commitment model are also available from the authors upon request. They point in the same direction.

[24] Here, we assume that if indifferent, the insuree buys from the informed (or big data) insurer.

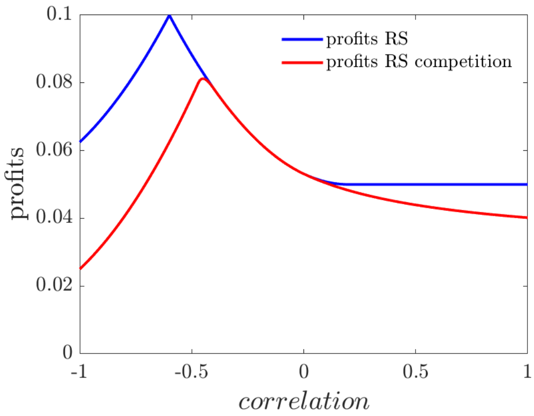


Figure 7: profit as a function of ρ with competition

in total information produced at the optimum due to the introduction of competition. This points to both a conceptual and practical challenge as to how to produce the most amount of information in the market without hurting the consumer surplus.

6.2 A regulation for transperancy

Given that the insurer wants to mislead the insuree, the government or a regulator might want to create a policy that forces the insurer to reveal all information it has. In our environment, this is equivalent to the insurer disclosing the correlation ρ. If the insurer is forced to reveal ρ, the environment is akin to Rothschild-Stiglitz environment, since now only the insuree has private information. Therefore, the insurer will implement the optimal Rothschild-Stiglittz menu. How does this policy affect the outcomes?

As we pointed out before in Proposition 2, in the information design model, the insurer divides the correlations in two areas: one in which there is complete pooling at the correlation ρ^* and another in which the insurer fully discloses its information. In the complete pooling area, the insurer offers the RS optimal menu at the correlation ρ^*. In this contract, the insurer does not offer any information rent to the insuree, since at that correlation, the insuree does not have any private information. Therefore, for those correlations, the insuree benefits from any policy that forces the insurer to reveal more information. Meanwhile, in the full revelation area, the insurer offers the optimal RS contract, so that the insuree is not affected by the government intervention. Therefore, the insuree is strictly better off for some correlations, and he is never worse off. Moreover, as seen before in the case with competition, both total surplus and the insurer's profit decrease strictly from the intervention in comparison to the standard information design problem.

7 Final remarks

A big debate in ensuing right now on the merits of technological advancements in data documentation and processing. Foregrounding these issues, in the summer of 2019, the New York Times carried a series of articles under the rubric of *The Privacy Project*.[25] One of the key topics of discussion therein was the impact of big data and AI on the insurance industry. This paper is an attempt to mainstream these discussions in the modeling choices made by classical economic theory in formalizing the key ideas in insurance markets.

Traditionally mechanism design models of insurance assume that the agent (or insuree) has some private information about the probability of incurring a loss or meeting with an accident. This results in the proverbial rent-versus-efficiency trade-off wherein the principal (or insurer) gives up on efficiency and provides information rents in order to separate the high risk from the low risk agents. We depart from this standard model in one crucial way— we make the state of world that parametrizes the loss to be two dimensional, and allow the agent to posses information about one of these dimensions and the principal to know the statistical correlation between the two dimensions. This creates an informed principal problem where the principal too has private information.

Private statistical information on the side of the insurer introduces a novel trade-off between obfuscation and price discrimination, in addition to the usual rent versus efficiency in standard screening contracts. The insurer wants to price discriminate using her private information dimension but is also wary that fine-tuning the contract too much to the details of the environment will allow the insuree to infer that information. This latter desire to maintain a obfuscation pulls against the desire to price discriminate.

In the standard framework in which the agent is Bayesian sophisticated, the information design model produces a bang-bang solution: the principal pools all types at the best possible obfuscation point, to the extent allowed by the Bayes plausibility condition and does full disclosure in the rest of the type space. There are significant gains from inverse selection in this scenario. Most of these gains are washed away if the principal cannot commit to the information structure, and typically one or two contracts are offered at the optimum—the ability of the agent to do Bayesian inference correctly protects it from incurring too much price discrimination.

While the results for the Bayesian sophisticated model can be considered as a theoretical benchmark, at the other extreme, if the agent is gullible, there are significant gains to be made from big data by exploiting the agent's limited inference capacities. This provides a foundation of sorts for both the rise of data markets and the returns to consumer activism whereby implications of data disclosure and its deployment by sellers can be better understood. Finally, putting this data in public domain along with an understanding of how to interpret this information can benefit consumers, and so will competition by endogenously limiting the extent to which the big data can be used against the consumers.

The ideas developed here can potentially be applied to contexts other than insurance. For example, in credit markets, owing to big data and AI, the credit issuing agency may also have some statistical information about the credit worthiness of a client, in addition to the client knowing some hard information about his financial circumstances. Finally, aggregating across multiple

[25] See www.nytimes.com/interactive/2019/opinion/internet-privacy-project.html.

principal-agent interactions. greater statistical information on the side of the principal may encourage more market concentration, of the kind we see in the tech-industry these days. Endogenizing data collection and market size is a promising question for future work.

8 Appendix

The appendix is divided into four parts.

8.1 Proof for benchmark model in Section 3

Proof of Proposition 1. Since ρ is common knowledge, the mapping r is redundant and the insurer's profit for a specific ρ is given by

$$\pi(\rho) = q_1(p_\rho(a) - \mu_\rho(a)x_\rho(a)) + (1-q_1)(p_\rho(b) - \mu_\rho(b)x_\rho(b)) \tag{5}$$

The insurer's problem is to choose a contract $c_\rho = \{p_\rho(\theta_1), x_\rho(\theta_1)\}_{\theta_1=a,b}$ to maximize $\pi(\rho)$ to subject to incentive feasibility. Using Equation (1), the constraints can be written as:

$$\mu_\rho(\theta_1)x_\rho(\theta_1) - \tfrac{\eta}{2}(1-x_\rho(\theta_1))^2 - p_\rho(\theta_1) \geqslant \mu_\rho(\theta_1)x_\rho(\theta_1') - \tfrac{\eta}{2}(1-x_\rho(\theta_1'))^2 - p_\rho(\theta') \quad \forall \theta_1, \theta_1' \in \{a,b\} \quad IC_{\theta_1-\theta_1'}$$
$$\mu_\rho(\theta_1)x_\rho(\theta_1) - \tfrac{\eta}{2}(1-x_\rho(\theta_1))^2 - p_\rho(\theta_1) \geqslant -\tfrac{\eta}{2} \quad \forall \theta_1 \in \{a,b\} \quad IR_{\theta_1}$$

Let ρ^* be the correlation for which $\mu_\rho(b) = \mu_\rho(a)$. Thus, when $\rho = \rho^*$, there is no asymmetric information, and the principal/insurer does not need to provide an information rent to any of the types. She can simply maximize efficiency by offering a unique pooling contract of full insurance $x_{\rho^*}(b) = x_{\rho^*}(a) = 1$ and bind the IR constraints to extract expected surplus.

Suppose that $\mu_\rho(b) > \mu_\rho(a)$, that is, $\rho > \rho^*$. Then, we are in the standard Rothschild and Stiglitz [1976] setup where $\theta_1 = b$ is the "high" or H type and $\theta_1 = a$ is the "low" or L type. It is standard practice (see for example Laffont and Martimort [2009]) to show that in this case IC_H and IR_L bind, and IC_L and IR_H are slack. Let λ be the multiplier on IC_H and δ the multiplier on IR_L. Th following FOCs that characterize an interior solution:

$$\begin{aligned}
[p_\rho(L)]:&\quad q_1 - \delta + \lambda = 0\\
[x_\rho(L)]:&\quad -\mu_\rho(L)q_1 + \delta\mu_\rho(L)q + \sigma(1-x_\rho(L))\delta - \lambda\mu_\rho(H) - \lambda\eta(1-x_\rho(L)) = 0\\
[p_\rho(H)]:&\quad (1-q_1) - \lambda = 0\\
[x_\rho(H)]:&\quad -(1-q_1)\mu_\rho(H) + \lambda\mu_\rho(H) + \eta(1-x_\rho(H))\lambda = 0.
\end{aligned}$$

From the first and third conditions it can be concluded that $\lambda = (1-q_1)$ and $\delta = 1$. Using these values it is straightforward to see that

$$x_\rho(H) = 1 \text{ and } x_\rho(L) = 1 - \frac{1-q_1}{\eta q_1}(\mu_\rho(H) - \mu_\rho(L)) < 1.$$

In case of a corner solution, $x_\rho(H) = 1$ and $x_\rho(L) = 0$.

Here of course $H = b$ and $L = a$. An analogous argument shows the result for the case in which $\mu_\rho(b) < \mu_\rho(a)$, that is, when $\rho < \rho^*$, and thus $H = a$ and $L = b$. □

8.2 Proofs for the information design problem in Section 4.1.

Proof of Proposition 2. The proof is divided in to seven steps, starting with two preliminary results. In what follows, we will abuse notation slightly and refer to $f(\rho, m)$ as the joint distribution generated by the information structure over primitive types and messages, and $f(m|\rho)$ as the corresponding conditional distribution.

Step 1. Two intermediate results are first established. First, we show that *there cannot be a non-zero mass of posteriors both to the left of ρ^* and to the right of it.* This result almost immediately implies that the insurer wants to pool as much probability as possible in messages that generate the posterior belief ρ^*. Second, we show that *the function π^{RS} is strictly convex* whenever the solution to the Rothschild-Stiglitz problem is interior and is constant otherwise.

Suppose, first, that the insurer sends messages that generates posterior beliefs for the insuree with non-zero measure on both sides of ρ^*. Call these the original messages. Notice that, by definition of ρ^*, at any of those posterior beliefs, the insurer obtains a profit that is smaller than $\pi^{RS}(\rho^*)$. Now, it is always possible for the insurer to reassign some probability from these messages to a new message m' such that $\mathbb{E}[\rho \mid m'] = \rho^*$. Moreover, since ρ^* in between these posterior beliefs, the reassignment can be done without changing the posterior beliefs generated from the original messages. As a consequence, the insurer strictly prefer this new information structure. This process can clearly be repeated as long as mass on at least one side of ρ^* has been exhausted, that is, as long as there is positive measure of beliefs in both sides of ρ^*. Therefore, we conclude that posterior beliefs are located at ρ^* and on one side of this point.

Second, suppose that $\mu_\rho(b) > \mu_\rho(a)$, that is, $\rho > \rho^*$. If we plug in the solution to Rothschild-Stiglitz problem in the insurer's profit function, we obtain that at an interior solution the profit function is equal to

$$\frac{\eta}{2} - (1 - q_1)(\mu_\rho(b) - \mu_\rho(a)) + \frac{(1 - q_1)^2}{2\eta q_1}(\mu_\rho(b) - \mu_\rho(a))^2.$$

This function is clearly strictly convex with respect to $(\mu_\rho(b) - \mu_\rho(a))$ and this probability gap is increasing in ρ. Therefore, $\pi^{RS}(\rho^*)$ is strictly convex if the solution to Rothschild-Stiglitz problem is interior. Further, when the solution to Rothschild-Stiglitz problem is a corner solution, the coverage that is offered to the insuree is independent of ρ and the insurer appropriates all the surplus that is created by the mechanism, which is constant.

Step 2. Now, we prove the three cases in the statement of the proposition. Consider first the case in which $\rho^* = \mathbb{E}[\rho]$. Since there cannot be a positive measure of posteriors on both sides of ρ^*, Bayes' plausibility immediately implies that the buyer's posterior equals ρ^* with probability one.

Step 3. Suppose now that $\mathbb{E}[\rho] > \rho^*$. Since there can be a positive mass only at ρ^* and on one side of ρ^*, Bayes plausibility immediately implies that this positive mass must be towards the right of ρ^*. So, no messages can generates posterior expectations that are smaller than ρ^*. Keeping this in mind, we want to characterize the complete pooling and full separation regions as stated in the proposition. Define ρ^{**} as the value that solves $\rho^* = \mathbb{E}[\rho \mid \rho \leq \rho^{**}]$. This threshold, we will show, demarcates the two regions. In particular, we will show that there is complete pooling for all $\rho \leqslant \rho^*$ exactly at the posterior expectation ρ^* and full disclosure for all $\rho > \rho^{**}$.

First, we claim that for $\rho \leqslant \rho^*$ the insurer sends a message that generates a posterior expectation for the insuree that is exactly equal to ρ^*. Suppose instead that there is a set of messages $M \subset \mathcal{M}$ such that $\int_M \int_{\underline{\rho}}^{\rho^*} f(m \mid \rho')f(\rho')d\rho' dm > 0$ and that for each of those messages the buyer's posterior is larger than ρ^* (remember it cannot be less than ρ^*).

Notice that by definition of ρ^*, $\pi^{RS}(\rho) < \pi^{RS}(\rho^*)$ for any $\rho > \rho^*$. And we just argued above that there has to be a set of messages $M' \subset \mathcal{M}$ such that $\int_{M'} \int_{\rho^*}^{\bar{\rho}} f(m \mid \rho')f(\rho')d\rho' dm > 0$ and the posterior expectation for each of those messages from the perspective of the insuree is larger than ρ^*. The insurer can thus be strictly better off by assigning as much probability as possible from the set of messages $\mathcal{M}$ and $\mathcal{M}'$ into a new message m'' that generates the posterior ρ^*, giving us a contradiction. So, it must be that with probability one, for $\rho \leqslant \rho^*$ the insurer sends a message that generates a posterior expectation for the insuree that exactly equals ρ^*.

Step 4a. Second, we claim that for $\rho^* < \rho \leqslant \rho^{**}$, the insurer also sends a message that generates a posterior expectation for the insuree that is exactly equal to ρ^*. Suppose by contradiction that exists a set of messages $\mathcal{M}$ such that $\int_{\mathcal{M}} \int_{\rho^*}^{\rho^{**}} f(m \mid \rho)f(\rho)d\rho dm > 0$ and that for each of those messages the buyer's posterior expectation is larger than ρ^*. Recollect that by definition, ρ^{**} satisfies the identity $\rho^* = \mathbb{E}[\rho \mid \rho \leq \rho^{**}]$. And all types $\rho \leqslant \rho^*$ put weight exclusively on ρ^*. So, there must be another set of messages $\mathcal{M}'$ such that $\int_{\mathcal{M}'} \int_{\rho^{**}}^{\bar{\rho}} f(m' \mid \rho')f(\rho')d\rho' dm' > 0$ and that for each of those messages the insuree's posterior is equal to ρ^*. If not, then the Bayes' plausibility condition that rests of the aforementioned identity will be violated. Using this information, we will create an alternate informtion structure for the insurer that does strictly better.

Step 4b . Take a message $m' \in M'$. Consider moving a ϵ proportion of the density $f(\rho', m')$ for each $\rho' \in [\rho^{**}, \bar{\rho}]$ to some message $m \in M$ and moving a δ proportion of the density $f(\rho, m)$ for each $\rho \in [\rho^*, \rho^{**}]$ to the message m'. When δ is chosen appropriately, we will show that the expected correlation after observing the message m' has not changed, $\mathbb{E}[\rho \mid m'] = \rho^*$, and that the conditional expected correlation after observing the message m has increased.

To prove this, let $\delta = \epsilon \frac{E(\rho',m') - \rho^* G(\rho',m')}{E(\rho,m) - \rho^* G(\rho,m)}$, with $E(\rho', m') = \int_{\rho^{**}}^{\bar{\rho}} \rho' f(\rho', m')d\rho'$, $G(\rho', m') = \int_{\rho^{**}}^{\bar{\rho}} f(\rho', m')d\rho'$, $E(\rho, m) = \int_{\rho^*}^{\rho^{**}} \rho f(\rho, m)d\rho$ and $G(\rho, m) = \int_{\rho^*}^{\rho^{**}} f(\rho, m)d\rho$. Notice that the conditional expectation $\mathbb{E}[\rho' \mid m'] = \rho^*$ can be written as

$$\frac{A' + \epsilon E(\rho', m')}{B' + \epsilon G(\rho', m')}$$

with $A' = \int_{\underline{\rho}}^{\bar{\rho}} \rho' f(\rho', m')d\rho' - \epsilon E(\rho', m')$ and $B' = \int_{\underline{\rho}}^{\bar{\rho}} f(\rho', m')d\rho' - \epsilon G(\rho', m')$. Therefore,

$$\begin{aligned}
&\rho^* = \frac{A' + \epsilon E(\rho', m')}{B' + \epsilon G(\rho', m')} \\
\Leftrightarrow & B'\rho^* = A' + \epsilon E(\rho', m') - \epsilon \rho^* G(\rho', m') \\
\Leftrightarrow & B'\rho^* = A' + \delta \frac{E(\rho, m) - \rho^* G(\rho, m)}{E(\rho', m') - \rho^* G(\rho', m')}(E(\rho', m') - \rho^* G(\rho', m')) \\
\Leftrightarrow & B'\rho^* + \delta \rho^* G(\rho, m) = A' + \delta E(\rho, m) \\
\Leftrightarrow & \rho^* = \frac{A' + \delta E(\rho, m)}{B' + \delta G(\rho, m)}
\end{aligned}$$

Notice that, by definition, $E(\rho', m') > E(\rho, m)$. Further, $\frac{E(\rho', m')}{G(\rho', m')} = \int_{\rho^{**}}^{\bar{\rho}} \rho' f(\rho' \mid m')d\rho' > \int_{\rho^*}^{\rho^{**}} \rho f(\rho \mid m)d\rho = \frac{E(\rho, m)}{G(\rho, m)}$. Using the definition of δ we conclude that $\delta G(\rho, m) > \epsilon G(\rho', m')$, and by using the equality above we conclude that $\delta E(\rho, m) > \epsilon E(\rho', m')$. Therefore, the insurer assigns a larger density to posterior ρ^* after the change in the information environment

Now, let $\tilde{\rho}$ be the initial posterior after observing the message m. This posterior can be written as $\tilde{\rho} = \frac{A + \delta E(\rho, m)}{B + \delta G(\rho, m)}$ with $A = \int_{\underline{\rho}}^{\bar{\rho}} \rho f(\rho, m)d\rho - \delta E(\rho, m)$ and $B = \int_{\underline{\rho}}^{\bar{\rho}} f(\rho, m)d\rho - \delta G(\rho, m)$. We want to show that $\frac{A + \epsilon E(\rho', m')}{B + \epsilon G(\rho', m')} > \tilde{\rho}$. Consider the functions $s(\rho) = \delta E(\rho, m) - \delta \rho G(\rho, m)$ and $r(\rho) = \epsilon E(\rho', m') - \delta \rho G(\rho', m')$. By definition of δ, $s(\rho^*) = r(\rho^*)$ and $s'(\rho) = -\delta G(\rho, m) < -\epsilon G(\rho', m') = r'(\rho)$. As $\tilde{\rho} > \rho^*$, the inequality above implies that $s(\tilde{\rho}) < r(\tilde{\rho})$. Therefore,

$$\begin{aligned}
&\tilde{\rho} = \frac{A + \delta E(\rho, m)}{B + \delta G(\rho, m)} \\
\Leftrightarrow & B\tilde{\rho} = A + \delta E(\rho, m) - \delta \tilde{\rho} G(\rho, m) \\
\Rightarrow & B\tilde{\rho} < A + \epsilon(E(\rho', m') - \tilde{\rho} G(\rho', m')) \\
\Leftrightarrow & \tilde{\rho} < \frac{A + \delta E(\rho', m')}{B + \delta G(\rho', m')}
\end{aligned}$$

where in the third line, we use that $s(\tilde{\rho}) < r(\tilde{\rho})$.

If we repeat this step for each $m' \in M'$ we will end up with an information design that generates the insuree's posterior ρ^* with higher probability, and splits the posterior $\tilde{\rho}$ generated by a message $m \in M$ between ρ^* and a posterior larger than $\tilde{\rho}$. Convexity implies that the insurer is better off if he repeats this process until $\int_{M'} \int_{\rho^{**}}^{\bar{\rho}} f(m' \mid \rho') f(\rho')d\rho' dm' = 0$. Thus, we conclude that the insurer is going to pool all correlations lower or equal to ρ^{**} in the posterior ρ^*.

Step 6. Finally, convexity implies that for $\rho > \rho^{**}$ the insurer prefers weakly to report truthfully the correlation. The insurer's preference is strict if the solution to RS problem for that correlation is interior.

Step 7. The case in which $\mathbb{E}[\rho] < \rho^*$ is analogous, repeated steps 3 to 6 making appropriate and symmetric changes.

□

Proof of Proposition 3. Suppose $\frac{G\rho)}{G(\tilde{\rho})} \leq \frac{F(\rho)}{F(\tilde{\rho})}$ for all $\tilde{\rho} \in [\underline{\rho}, \bar{\rho}]$ and $\rho \leq \tilde{\rho}$. We first show that $\mathbb{E}_G[\rho \mid \rho \leq \rho'] = \int_{\underline{\rho}}^{\rho'} \rho \frac{g(\rho)}{G(\rho')} d\rho > \int_{\underline{\rho}}^{\rho'} \rho \frac{f(\rho)}{F(\rho')} d\rho = \mathbb{E}_F[\rho \mid \rho \leq \rho']$ for any ρ'. We have that

$$\begin{aligned}
\int_{\underline{\rho}}^{\rho'} \rho \frac{g(\rho)}{G(\rho')} d\rho &= \rho \frac{G(\rho)}{G(\rho')} \Big|_{\underline{\rho}}^{\rho'} - \int_{\underline{\rho}}^{\rho'} \frac{G(\rho)}{G(\rho')} d\rho \\
&= \rho' - \int_{\underline{\rho}}^{\rho'} \frac{G(\rho)}{G(\rho')} d\rho \\
&\geq \rho' - \int_{\underline{\rho}}^{\rho'} \frac{F(\rho)}{F(\rho')} d\rho \\
&= \int_{\underline{\rho}}^{\rho'} \rho \frac{f(\rho)}{F(\rho')} d\rho
\end{aligned}$$

Consider the case in which $\mathbb{E}_F[\rho] > \rho^*$. The result above implies that $\mathbb{E}_G[\rho] > \rho^*$. Then we are in the second case of Proposition 2. From the definition of ρ^{**} in the proof of Proposition 2 and the result above, we conclude immediately that $\rho_G^{**} \leq \rho_F^{**}$. Therefore, the set of correlations for which the seller discloses information truthfully is smaller when the distribution is G rather than F.

An analogous argument shows the result for the case in which $\mathbb{E}_G[\rho] < \rho^*$

□

8.3 Formal description, results and proofs for the limited commitment model in Section 4.2

Here we presented a formal treatment of the limited commitment model outlined and discussed intuitively in Section 4.2. Recollect that the optimization problem reads as follows:

$$(\mathcal{P} - LC): \quad \max_{r,C} \Pi \text{ s.t. } IC_\rho, IC_{\theta_1}, IR_{\theta_1}.$$

where the only difference from the information design problem is the addition of the incentive constraint on the side of the principal, IC_ρ. We start with the standard Myersonian characterization of the principal's incentive compatibility constraint.

Characterizing incentive compatibility

Lemma 1. *IC_ρ holds if and only if π satisfies the following*

1. *envelope characterization of local incentives:*

$$\frac{\partial \pi(\rho; \hat{\rho})}{\partial \rho}\Big|_{\hat{\rho}=\rho} = \sigma x_{r(\rho)}(L) \cdot (\mu_{ab} - \mu_{aa}) - \sigma x_{r(\rho)}(b) \cdot (\mu_{bb} - \mu_{ba}) \equiv c(\rho), \text{ and,} \quad (6)$$

2. *convexity: $\pi(\rho)$ is convex in ρ.*

Proof. Part two is a standard property of value functions that satisfy incentive compatibility on a continuous type space (see, for example, Börgers [2015], Chapter 3). We show here the exact functional form of the envelope characterization stated in Equation (6). Start with Equation (4), i.e. assuming truthteling by the insuree, the profit function from (mis)reporting $\hat{\rho}$ is given by

$$\pi(\rho;\hat{\rho}) = q_1\left[p_{r(\hat{\rho})}(a) - \mu_\rho(a)x_{r(\hat{\rho})}(a)\right] + (1-q_1)\left[p_{r(\hat{\rho})}(b) - \mu_\rho(b)x_{r(\hat{\rho})}(b)\right]$$

where the only terms that are a function of ρ are

$$\mu_\rho(a) = (q_2 + \rho\sigma/q_1)\mu_{aa} + ((1-q_2) - \rho\sigma/q_1)\mu_{ab}, \text{ and}$$
$$\mu_\rho(b) = (q_2 - \rho\sigma/(1-q_1))\mu_{ba} + (1 - q_2 + \rho\sigma/(1-q_1))\mu_{bb}$$

Taking a derivative with respect to ρ, then gives us:

$$\frac{\partial\pi(\rho;\hat{\rho})}{\partial\rho} = -\sigma x_{r(\hat{\rho})}(a)(\mu_{aa} - \mu_{ab}) - \sigma x_{r(\hat{\rho})}(b)(-\mu_{ba} + \mu_{bb})$$

and, substituting $\hat{\rho} = \rho$ delivers Equation (6). □

By fixing r, we fix M, which partitions type space of possible correlations, $[\underline{\rho}, \overline{\rho}]$. Hence we also fix the number of contracts offered at the optimum, $|C| = |M|$. Now, for a given r, Lemma 1 tells us two things. First, the slope of the profit function can be written as

$$c(\rho) = k_a\phi_a(\rho) - k_b\phi_b(\rho),$$

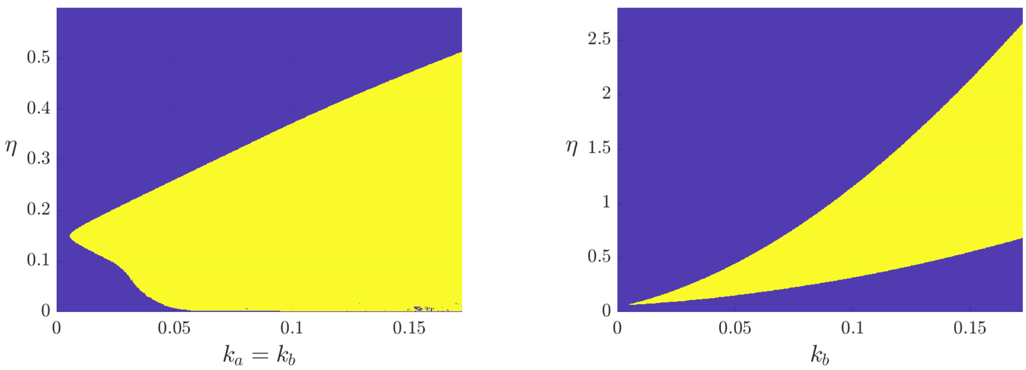
where k_a and k_b are positive constants, and $\phi_a(\rho) = x_{r(\rho)}(a)$ and $\phi_b(\rho) = x_{r(\rho)}(b)$ are the coverages chosen for $\theta_1 = a$ and $\theta_1 = b$, as a function of the partition of $\mathcal{M}$ in which ρ falls. And, second, by convexity of π, that $c(\rho)$ must be non-decreasing. These two together put restrictions on what coverages/allocations are feasible, specifically they limit the extent of price discrimination that the insurer can employ even for a fixed number of contracts.

Finiteness and then fewness of the optimal contract

The typical approach taken in mechanism design is to ignore the convexity constraints, solve the relaxed problem using only the envelope condition, and invoke a regularity condition such as the monotone hazard rate. But this problem is not standard in at least three ways: (i) the "policy function" is multidimensional, there are two allocation rules in the envelope condition, viz. ϕ_a and ϕ_b, (ii) these functions in turn solve another downstream screening problem for the agent, and (iii) the mechanism still has to jointly choose r and C at the optimum.

All of the aforementioned constraint the contract space in non-trivial ways, and the convexity constraint cannot be ignored generically. We show that an optimal contract must in fact be finite:

Lemma 2. *The optimal mechanism has a finite number of messages and contracts:* $|\mathcal{M}| = |C| = \kappa$ *for some* $\kappa \in \mathbb{N}$.



(a) case $k_a = k_b$

(b) fixing k_a, varying k_b

Figure 8: Splitting the parameters space into two regions: $|\mathcal{M}| = |C| = 1$ and $|\mathcal{M}| = |C| > 1$.

To further highlight the fewness of contracts at the optimum, consider the limiting case when effective risk in the model become large. Recollect that $\eta = \gamma \mathbb{V}[l]$, where γ is the risk aversion parameter, and $\mathbb{V}[l]$ is the variance of loss. Thus, η is a sufficient statistic of risk in our setup. We have:

Lemma 3. *As $\eta \to \infty$, insurer finds it optimal to offer one message and contract, $|\mathcal{M}| = |C| = 1$.*

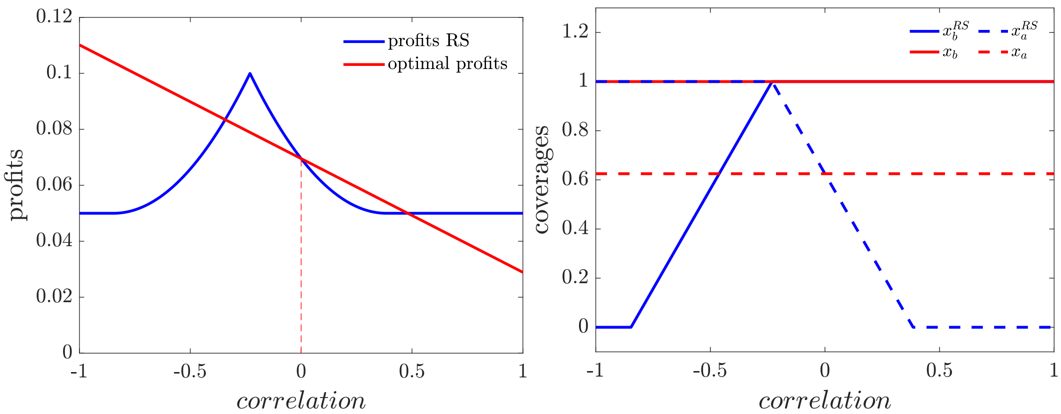
When the insurer sets $|\mathcal{M}| = |C| = 1$, there is complete pooling across ρ. Thus, she simply does not use her informational advantage towards price discrimination, opting rather to maintain the ex ante belief gap. Thus, the power of Bayesian inference essentially compels the insurer to resolve the trade-off between obfuscation and price discrimination completely in favor of the former force, at least in the two limits specified in Lemma 3. Lemmata 2 and 3 together prove Proposition 4.

To give a richer flavor to the fewness of contracts, we also report some numerical results for intermediate values of η. Recollect that the number of contracts offered at the optimum depends on the slope of the profit function: $c(\rho) = k_a\phi_a(\rho) - k_b\phi_b(\rho)$, which in turn depends on the primitives k_a and k_b, and the allocations $\phi_a(\rho)$ and $\phi_b(\rho)$. The allocations are of course driven by the extend of risk and uncertainty in the environment, viz η. Hence, to understand the structure of optimal contracts, we split parametric space along these two dimensions.

Figure 8 documents when the optimal contract features complete pooling, $|\mathcal{M}| = |C| = 1$, and when it features some price discrimination along ρ, that is $|\mathcal{M}| = |C| > 1$. In fact, in the numerical simulation we allow the program to accept up to three partitions and it selects either one (in the dark/purple region) or two (in the light/yellow region) but never three. This further provides credence to the claim that the *number of contracts offered at the optimum with Bayesian sophisticated agents is small.*

Optimal one and two partition contracts

If the optimal number of partitions turns out to be one, it is fairly intuitive to conclude that the coverages offered would be same as those offered in the Rothschild and Stiglitz [1976] model at the



(a) profit (b) coverage and its benchmark counterpart

Figure 9: Optimal contract features complete pooling

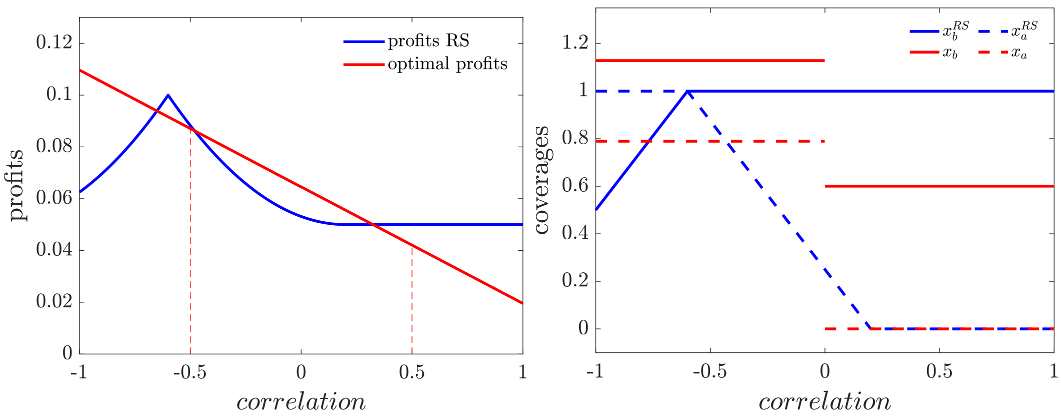
ex ante expected correlation, and the optimal profit too will be equal to the optimal profit at that correlation. In the interest of space, we illustrate the result in a graph, the formal statement and proof are available from authors upon request.

Figure 9 plots the optimal profit and coverages for this case. The coverages are simply straight horizontal lines for the insurer is not using any of her private information about ρ and instead offers a completely pooling contract along ρ. As in the benchmark model, the "high" risk insuree (which is type $\theta_1 = b$ in the figure) is given full insurance and the "low" risk insuree is given partial insurance. The profit function is a straight downward sloping line since the allocations are fixed, and π is linear in ρ. The dotted vertical line captures the expected correlation at which point the red straight line and benchmark blue curve intersect.[26]

Next, we consider the case where the optimal number of partitions is two. In this case, the type space of correlations is split into two intervals, say $\mathbb{I}_1$ and $\mathbb{I}_2$. The coverages in each interval are evaluated using the expected correlation in those intervals while ensuring that the insuree's incentive constraint is satisfied between reporting interval $\mathbb{I}_1$ or $\mathbb{I}_2$ and within each interval, the insurer's incentive constraint is satisfied in reporting $\theta_1 = b$ or a. Again we illustrate the main result through graphs and formal statements and proofs are available from authors upon request.

Figure 10 plots the optimal profit and coverages when the optimal number of partitions is two. Each partition corresponds to two coverages, one for each insuree type, which gives the profit function its slope. The first result is that optimality forces both these slopes to be the same. This follows from the intuition given in Section 4.2 that without imposing the convexity constraint the optimal profit line has a concave link. So the highest profit line that satisfies convexity is then simply the straight line which equates the slope of the profit function along the two partitions, as shown in Figure 10a. The second result is simply that if the insuree is employing two partitions at the optimum than the profit must be greater than expected profit in the benchmark model since the latter can always be attained by offering a completely pooling contract.

[26]We plot the profit and the coverages of the benchmark model simultaneously to help motivate the impact of the privacy of statistical information on the side of the insurer, which separates our model form (most of) the literature on insurance markets.



(a) profit (b) coverage and its benchmark counterpart

Figure 10: Optimal contract features two partitions

The third result documents that the left partition where $\theta_1 = a$ is the "high" risk type, $\theta_1 = b$ is overinsured. But incentive constraints force the allocation to always satisfy $x_1(b) \geqslant x_1(a)$. So, in the right partition, $\theta_1 = b$, which is now the "high" risk type, is offered under insurance, and $\theta_1 = a$ is forced out of the market with no coverage. See Figure 10b.

Summary

To summarize, when the insuree is Bayesian sophisticated and the insurer lacks commitment to the information disclosure policy, the total number of contracts offered at the optimum is small, often at most two. This illustrates the fact the trade-off between obfuscation and price discrimination is resolved mostly in favor of the former. When the optimal contract features complete pooling across ρ, the contract corresponding to the benchmark model at the expected correlation is offered and the insurer does not use her informational advantage at all. When the optimal number of contracts offered is two, we see both over and under insurance at the optimum—the insurer tried to mislead the insuree to the extent possible by the Bayesian rationality.

Proofs for Lemmata 2 and 3

Proof of Lemma 2. To prove this result, we will first introduce the concept of interim profit for the insurer, and state and prove a lemma characterizing it. Define the function $\hat{\pi}(\rho, c)$ to be the maximum profit that an insurer can obtain in the subgame in which both parties believe that the correlation is ρ, and the slope of the profit function as defined by Equation (6) is equal to c, that is $c(\rho) = c$.

The analysis here is stated for the case when ρ is common knowledge in the subgame, that is, r is the identity mapping, $r(\rho) = \rho \ \forall \ \rho$. Since profits ($\pi$) and utility ($u$) are both linear in ρ, we will later replace ρ with the expectation of the partition in which it lies, and all results stated here would carry through.

The following lemma establishes that the interim profit function, $\hat{\pi}(\rho, c)$ is (i) single peaked with respect to ρ, with a peak at ρ^*, (ii) it is convex with respect to ρ both to the right and the left of ρ^*; and (iii) it is strictly concave with respect to c. The next result states some other key properties of the interim profit function.

Lemma 4. *Let ρ^* solve $\mu_\rho(b) = \mu_\rho(a)$ and fix $\rho \geq \rho^*$. Then:*

1. *$\exists$ ρ_1 and ρ_2 with $\bar{\rho} \geq \rho_2 \geq \rho_1 \geq \rho^*$ such that*

   (a) for $\rho \in [\rho^*, \rho_1]$, $\hat{\pi}(\cdot, c)$ is linear and strictly decreasing in ρ;

   (b) for $\rho \in [\rho_1, \rho_2]$, $\hat{\pi}(\cdot, c)$ is strictly convex and strictly decreasing in ρ;

   (c) for $\rho > \rho_2$, $\hat{\pi}(\cdot, c)$ is constant in ρ.

2. *The function $\hat{\pi}(\rho, \cdot)$ is strictly concave in c.*

Analogous characterization holds for $\rho < \rho^*$.

Proof. We prove the result in 5 steps.

Step 1. Fix $\rho \geq \rho^*$, so that $\mu_\rho(b) \geqslant \mu_\rho(a)$. To simplify notation, let $\mu_a = \mu_\rho(a)$, $\mu_b = \mu_\rho(b)$, $x_a = x_\rho(a)$, $x_b = x_\rho(b)$, $p_a = p_\rho(a)$ and $p_b = p_\rho(b)$. As in the main text, let $k_a = \sigma(\mu_{ab} - \mu_{aa})$ and $k_b = \sigma(\mu_{bb} - \mu_{ba})$. We prove part 1 first.

Step 2. In an interior solution, we obtain that:

$$\begin{aligned} x_a &= 1 - \tfrac{1-q_1}{\eta q_1}(\mu_b - \mu_a) + \tfrac{\beta}{\eta q_1} k_a \\ x_b &= 1 - \tfrac{\beta}{\eta(1-q_1)} k_b \end{aligned}$$

where $\beta = \frac{\eta q_1(1-q_1)(c-k_a+k_b)+(1-q_1)^2(\mu_b-\mu_a)k_a}{(1-q_1)k_a^2+q_1k_b^2}$ is the Lagrange multiplier for the convexity constraint. Further, substituting for x_a and x_b, the optimal profit is

$$\hat{\pi} = \frac{\eta}{2} - (1-q_1)(\mu_b - \mu_a) + \frac{(1-q_1)^2}{2\eta q_1}(\mu_b - \mu_a)^2 - \frac{(1-q_1)k_a^2 + q_1 k_b^2}{2\eta q_1(1-q_1)}\beta^2.$$

Since β depends on ρ only through $\mu_b - \mu_a$, π is quadratic in ρ. Its first derivative with respect to ρ is

$$(1-q)\left(\frac{\partial(\mu_b - \mu_a)}{\partial \rho}\right)\left(-1 + \frac{1-q_1}{\eta q_1}(\mu_b - \mu_a) - \frac{\beta k_a}{\eta q_1}\right) < 0,$$

since the partial derivative is positive and the last term has to be negative to guarantee that x_a is positive. Its second derivative with respect to ρ is given by,

$$\frac{(1-q_1)^2 k_b^2}{\eta((1-q_1)k_a^2 + q_1 k_b^2)}\left(\frac{\partial(\mu_b - \mu_a)}{\partial \rho}\right)^2 > 0.$$

Therefore, $\pi(\rho, c)$ is strictly decreasing and strictly convex with respect to ρ when the solution is interior. Next we will construct the bounds for interiority: ρ_1 and ρ_2.

Step 3. A corner solution in which the insurer sells only to $\theta_1 = b$ insuree occurs when x_a is negative, that is, when $\frac{\eta(q_1 k_b^2+(1-q_1)(c+k_b)k_a)}{(1-q_1)k_b^2} < \mu_b - \mu_a$. Since $\mu_b - \mu_a$ is increasing in ρ, this condition may hold only for large correlations. Let ρ_2 to be equal to the correlation that makes this condition to hold with equality if it is smaller than $\bar{\rho}$ and equal to $\bar{\rho}$, otherwise.

In such a corner it has to be that $x_b = \max\{\frac{-c}{k_b}, 0\}$. Using the constraint IR_b, we obtain that $(1-q_1)(p_b - \mu_b x_b) = (1-q_1)\max\{\frac{-\eta c}{2k_b}\left(2+\frac{c}{k_b}\right), 0\}$, so that the profit function is constant in ρ.

Step 4. Finally, to satisfy both IC constraints of the isnuree it has to be that $x_b \geqslant x_a$, but this might not be the case in the interior solution we characterized above. In particular for $\rho < \rho_2$ the constraint $x_b \geqslant x_a$ is not satisfied if

$$\begin{cases} \mu_b - \mu_a < \frac{\eta(c-k_a+k_b)((1-q_1)k_a+q_1k_b)}{-(1-q_1)k_b(-k_b+k_a)} & \text{if } -k_b+k_a > 0 \\ \mu_b - \mu_a > \frac{\eta(c-k_a+k_b)((1-q_1)k_a+q_1k_b)}{-(1-q_1)k_b(-k_b+k_a)} & \text{if } -k_b+k_a < 0. \end{cases}$$

Notice that in the first case the inequality is never true if $c-k_a+k_b < 0$ and in the second case it is never true if $c-k_a+k_b > 0$. In the domain in which the inequalities can be true, the correlation that makes the first inequality to hold with equality is smaller than ρ_2, and the correlation that makes the second inequality to hold with equality is larger than ρ_2. Then we define ρ_1 in the first case as the maximum of the correlation that makes the inequality to hold with equality and ρ^*, and in the second case we just define it as ρ^*.

For correlations in $[\rho^*, \rho_1]$, the insurer offers a unique package $x_b = x_a = \frac{c}{-k_b+k_a} > 0$ at the price that makes the constraint IR_a to hold with equality. This generates profits equal to

$$\hat{\pi} = \frac{\eta}{2} - \frac{\eta}{2}\left(1 - \frac{c}{k_a - k_b}\right)^2 - (1-q_1)(\mu_b - \mu_a)\frac{c}{k_a - k_b}.$$

Since the profits depend on ρ only trough $\mu_b - \mu_a$ and this dependence is linear, we conclude that the profits are linear with a slope $s_1 = -\frac{((1-q_1)k_a+q_1k_b)c}{q_1(k_a-k_b)} < 0$.

The second inequality is true only for correlations for which the constraint $x_a \geqslant 0$ binds as well, that is, in the solution to the problem without these constraints both x_a and x_b are negative. Therefore, since the b type is willing to pay more for insurance, the insurer sells only to him, and we are back to the case in step 3.

Step 5. To prove 2, we only need to take the second derivative of the profit functions with respect to c for all possible cases. In the corner solution with $x_a = 0$ and $x_b > 0$ we have

$$\frac{\partial^2 \hat{\pi}}{\partial c^2} = \frac{-(1-q_1)\eta}{k_b^2} < 0,$$

in the corner solution with $x_a = x_b > 0$ we have

$$\frac{\partial^2 \hat{\pi}}{\partial c^2} = \frac{-\eta}{(k_a - k_b)^2} < 0,$$

and in the interior solution with $x_b > x_a > 0$ we have that

$$\frac{\partial^2 \hat{\pi}}{\partial c^2} = \frac{-\eta q_1(1-q_1)}{(1-q_1)k_a^2 + q_1 k_b^2} < 0.$$

Therefore, the function $\hat{\pi}(\rho, \cdot)$ is strictly concave with respect to c. □

Now we proceed to prove the finiteness result in five steps. First, we argue that the optimal information disclosure policy is ordered, i.e., can be regarded as a partition. This automatically implies that the optimal mechanism has at most a countable number of partitions, and then we use the next four steps to gradually build towards the finiteness of the number of partitions.

Step 1. Define $r^{-1}(m)$ be the set of correlations for which is possible that the insurer sends the message m, that is, $r^{-1}(m) = \{\rho \mid \mathbb{P}(m \mid \rho) > 0\}$. We say that a message rule is *ordered* if for any m and $\hat{m}$ and any correlations $\rho \in r^{-1}(m)$ and $\hat{\rho} \in r^{-1}(\hat{m})$ either $\rho < \hat{\rho}$ or $\rho > \hat{\rho}$. Now, we show that it is without loss of generality for the insurer to use an *ordered* message rule. Note that this definition implies directly that *an ordered message rule is deterministic in the sense that given a correlation* ρ *the insurer always sends the same message.*

Suppose, in the optimal contract, the insurer uses a message rule that is not ordered. Then there exists correlations $\rho < \hat{\rho} < \rho'$ such that $\rho \in r^{-1}(m)$, $\rho' \in r^{-1}(m)$ and $\hat{\rho} \in r^{-1}(\hat{m})$. Since the contract has to satisfy the convexity constraint it has to be that the slope of the optimal profit satisfies the condition $c(\rho) \leq c(\hat{\rho}) \leq c(\rho')$. However, as for both ρ and ρ' the seller sends with positive probability the same message, the contract offered by the insurer for any of the two correlations must be the same, i.e., $c(\rho) = c(\rho')$. Both conditions imply that $c(\rho) = c(\hat{\rho}) = c(\rho') := c$, that is, the profit function is locally linear. Then the profit function is locally linear for correlations in the preimages $r^{-1}(m)$ and $r^{-1}(\hat{m})$ and any correlation $\hat{\hat{\rho}}$ such that there exist $\rho \in r^{-1}(m)$ and $\hat{\rho} \in r^{-1}(\hat{m})$ with $\rho < \hat{\hat{\rho}} < \hat{\rho}$ or $\hat{\rho} < \hat{\hat{\rho}} < \rho$.

Now, let $E_m = \mathbb{E}[\rho \mid \rho \in r^{-1}(m)]$ and $E_{\hat{m}} = \mathbb{E}[\hat{\rho} \mid \hat{\rho} \in r^{-1}(\hat{m})]$. The local linearity of the profit function π implies that $\pi(\rho) = \pi(E_m) + c(\rho - E_m) = \pi(E_{\hat{m}}) + c(\rho - E_{\hat{m}})$ for any $\rho \in r^{-1}(m) \bigcup r^{-1}(\hat{m})$. Therefore, for any reshuffle of correlations in that sets $r^{-1}(m)$ and $r^{-1}(\hat{m})$ that keeps E_m and $E_{\hat{m}}$ constant, the seller can attain the same profits when offering the same contract at those expected correlations. In particular, the insurer can always reshuffle the correlations in that sets $r^{-1}(m)$ and $r^{-1}(\hat{m})$ in a way that keeps E_m and $E_{\hat{m}}$ constant, and such that for any $\rho \in r^{-1}(m)$ and $\hat{\rho} \in r^{-1}(\hat{m})$, $\rho < \hat{\rho}$. Then, the insurer can obtain an ordered message rule that attains the same profits as the original message rule by repeating this process as many times as necessary.

Step 2. Suppose by contradiction that the profit function's slope takes an infinite number of different values. Fix $\epsilon_1 > 0$. As the slope take an infinite number of different values, there exists two different messages m_1 and m_2 with the following properties:

1. the preimages $r^{-1}(m_1)$ and $r^{-1}(m_2)$ are contiguous,

2. the contracts c_{m_1} and c_{m_2} generate different profits slopes $c_1 < c_2$ such that there are no others posted contracts that generate the same slopes, and

3. the insuree's expected correlations after observing those messages, $\rho_1 = E[\rho \mid r(\rho) = m_1] < \rho_2 = E[\rho \mid r(\rho) = m_2]$, are such that $\rho_2 - \rho_1 < \epsilon$.

Without loss of generality, assume that $\rho_1 > \rho^*$. By Proposition 1 we have that $c^{RS}(\rho_1) > c^{RS}(\rho_2)$. Further, Lemma 4 shows that the functions $\hat{\pi}(\rho_1, \cdot)$ and $\hat{\pi}(\rho_2, \cdot)$ are concave in c. Therefore, $c^{RS}(\rho_1)$ and $c^{RS}(\rho_2)$ correspond to the unique maximizers of these functions, respectively, because Rothschild-Stiglitz problem does not impose constraints on c. Therefore, for any $c < c^{RS}(\rho)$, $\hat{\pi}(\rho, \cdot)$ is increasing, and for any $c > c^{RS}(\rho)$, $\hat{\pi}(\rho, \cdot)$ is decreasing.

Step 3. Before delving into impossibility to infinite slopes, we show as an intermediate step that it cannot be that $c_1 < c^{RS}(\rho_1)$ and $c_2 > c^{RS}(\rho_2)$. If they were, c_1 would be in the increasing part of $\hat{\pi}(\rho_1, \cdot)$ and c_2 would be in the decreasing part of $\hat{\pi}(\rho_2, \cdot)$. Then, by slightly increasing c_1 and slightly decreasing c_2, the firm relaxes the feasibility constraints and is able to increase its profits (see Figure 11). Therefore, the original c_1 and c_2 cannot be optimal, a contradiction. Thus, we conclude that $c_1 \geq c^{RS}(\rho_1)$ or $c_2 \leq c^{RS}(\rho_2)$.

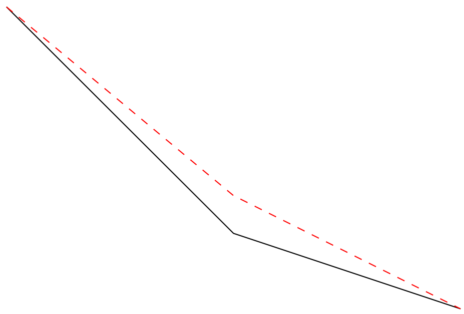


Figure 11: Improvement arguments for Step 3 of the proof.

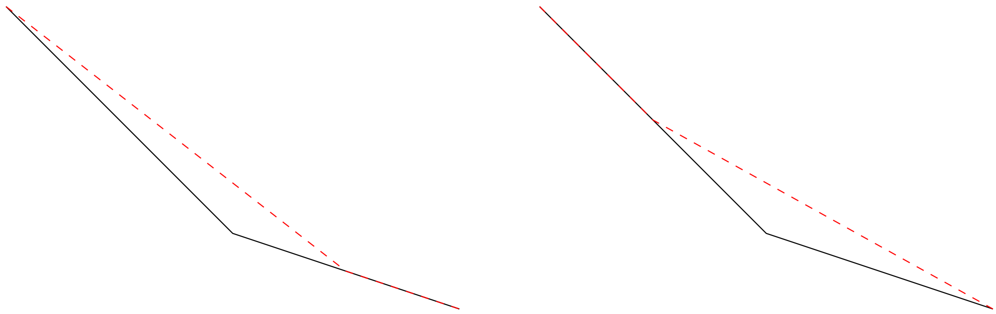
Step 4. Suppose first that $c_1 \geq c^{RS}(\rho_1)$. We show by contradiction that in this case whenever the slope takes an infinite number of different values, the resulting profit function cannot be continuous, contradicting the convexity constraint.

As $c_2 > c_1$ and $c^{RS}(\rho_1) > c^{RS}(\rho_2)$, the previous discussion implies that both c_1 and c_2 are located in the decreasing part of $\hat{\pi}(\rho_2, \cdot)$. Therefore, $\hat{\pi}(\rho_2, c_1) - \hat{\pi}(\rho_2, c_2) = \delta > 0$.

Let $\pi(.)$ denote the optimal profit function. By feasibility, it must be that $\pi(\rho_1) < \hat{\pi}(\rho_1, c_1)$ and $\pi(\rho_2) < \hat{\pi}(\rho_2, c_2)$. We argue that it has to be that $\pi(\rho_1) = \hat{\pi}(\rho_1, c_1)$ by showing perturbations that increase the seller's profits. We pick perturbations that do not change the function π outside the two partitions we are considering. There are two cases.

If both $\pi(\rho_1) < \hat{\pi}(\rho_1, c_1)$, and $\pi(\rho_2) < \hat{\pi}(\rho_2, c_2)$ then the insurer/seller can slightly increase c_1 and move ρ_2 to the right. Figure 12a shows that this allows the insurer to increase profits, and it is feasible because both inequalities are strict and all functions are continuous.

Instead, if $\pi(\rho_1) < \hat{\pi}(\rho_1, c_1)$ and $\pi(\rho_2) = \hat{\pi}(\rho_2, c_2)$, the insurer can slightly decrease c_2 and move ρ_1 to the left. As the initial c_2 is in the decreasing part of $\hat{\pi}(\rho_2, \cdot)$, this relaxes the constraint at the second partition and makes the perturbation feasible. Figure 12b depicts why this change benefits the firm.



(a) Increase c_1 and move ρ_2 to the right. (b) Decrease c_2 and move ρ_1 to the right.

Figure 12: Improvement arguments for Step 4 of the proof.

Let $\dot{\rho}$ be the threshold correlation between the two partitions. Fixed $\delta > 0$. By continuity of $\hat{\pi}(\rho, c)$ with respect to ρ there exists ϵ_2 such that if $|\ \rho_1 - \rho_2\ | < \epsilon_2$ then $|\ \hat{\pi}(\rho_1, c_1) - \hat{\pi}(\rho_2, c_1)\ | < \frac{\delta}{4}$. Take ϵ_2 small enough such that if $c_1 < 0$ then $\epsilon_2 < \frac{-\delta}{4c_1}$ and if $c_1 > 0$ then $\epsilon_2 < \frac{\delta}{4c_1}$.

Let $\epsilon = \min\{\epsilon_1, \epsilon_2\}$. Putting all our calculations together we obtain that

$$\begin{aligned} \pi(\rho_1) + c_1(\dot{\rho} - \rho_1) = \; & \hat{\pi}(\rho_1, c_1) + c_1(\dot{\rho} - \rho_1) \\ & > (\hat{\pi}(\rho_2, c_1) - \tfrac{\delta}{4}) - \tfrac{\delta}{4} \\ & = \hat{\pi}(\rho_2, c_2) + \delta - \tfrac{\delta}{2} \\ & > \hat{\pi}(\rho_2, c_2) + \tfrac{\delta}{4} + c_2(\rho_2 - \dot{\rho}) \\ & > \hat{\pi}(\rho_2, c_2) + c_2(\rho_2 - \dot{\rho}) \\ & \geq \pi(\rho_2) + c_2(\rho_2 - \dot{\rho}), \end{aligned}$$

but this implies that the pasting condition is not satisfied. This is a contradiction because the insurer's optimal profit function π is convex, and therefore it is continuous. Therefore, there cannot exist such elements of the partition.

An analogous argument leads to a contradiction in the case in which $c_2 \leq c^{RS}(\rho_2)$.

Step 5. Finally, we argue that in the optimal contract, there can be at most two elements of the partition that generate the same profit's slope c. Lemma 4 implies that $\hat{\pi}(\cdot, c)$ is convex. Therefore, a contract with three different messages that generate the same slope is trivially dominated by a contract with only two different messages that lead to the same slope.

So, the partition in the optimal contract can contain only a finite number of elements.

□

Proof of Lemma 3. Notice that $\mathbb{E}(\rho) > \rho^*$. From Proposition 1 we obtain that as $\eta \to \infty$, $x_b^{RS}(\rho) = 1$ and $x_a^{RS}(\rho) \to 1$ if $\rho > \rho^*$, and $x_a^{RS}(\rho) = 1$ and $x_b^{RS}(\rho) \to 1$ otherwise. That is, the contract that is offered in Rothschild-Stiglitz benchmark becomes independent of the correlation. Therefore, the insurer can approximate the optimal profits at each point by offering only one contract that consists of the Rothschild-Stiglitz menu at the ex-ante expected correlation. In the limit, when the insurer offers full insurance to both types, and obtains profit equal to $\frac{\eta}{2}$.

If instead the insurer decides to offer more than one contract, he has to face a non-vanishing cost that is imposed by the convexity constraint. Actually, from the proof of Lemma 4 can be observed that $\hat{\pi}(\rho, c)$ is always smaller and away from $\frac{\eta}{2}$ as long as $c \neq k_a - k_b$, the contract's slope that is generated by the full insurance contract. Therefore, in the limit the insurer cannot improve upon having just one contract. □

8.4 Formal statement and proof for Section 6.1

Note that the global maximum for the lower red curve in Figure 7 is at a new point to the right of ρ^*. We call this new peak as $\hat{\rho}^*$. We can now state the main result for the introduction of competition by an insurer that does not have access to big data.

Proposition 7. *Define $\hat{\rho}^{**}$ such that $\mathbb{E}[\rho \mid \rho < \hat{\rho}^{**}] = \hat{\rho}^*$, and let I_0 the insuree's information rent at the RS contract at ρ^e. Suppose $\rho^e > \rho^*$ and $\frac{2}{3}\left(\mu_{\hat{\rho}^{**}}(b) - \mu_{\hat{\rho}^{**}}(a)\right) > I_0$. Then we obtain the following results:*

- *The insurer discloses less information when there is competition. With competition, there is a pooling region in which the insurer sends the message $r(\rho) = \hat{\rho}^*$. This pooling region is a superset of the pooling region without competition.*
- *The insurer's ex-ante expected profit is lower with competition. The ex-ante expected profits in the pooling region $\hat{\rho}^*$ are strictly smaller with competition than without competition.*
- *Total surplus is the same in the pooling regions with and without competition. Total surplus is weakly larger with competition than without competition in the full disclosure region with competition.*
- *If $I_0 > \mu_{\hat{\rho}^*}(b) - \mu_{\rho^*}(b)$, ex-post consumer surplus is point-wise strictly larger with competition than without competition for correlations that belong to the pooling region without competition. Consumer surplus is always weakly larger for correlations in the full disclosure region with competition.*

An analogous result follows if $\mathbb{E}[\rho] < \rho^*$.

Proof. The proof is divided into 3 steps.

Step 1. We first characterize the insurer's optimal profits when ρ is common knowledge and there is competition, that is, the lower red curve in Figure 7.

Suppose $\mathbb{E}[\rho] > \rho^*$ and denote by I_0 type b insuree's information rent in the standard RS problem. Recall that since $\mathbb{E}[\rho] > \rho^*$ type a's information rent is zero. Given that both the insuree and the insurer know that the correlation is ρ, the insurer problem is

$$
\begin{aligned}
max_{x_\rho,p_\rho} & q_1(p_\rho(a) - \mu_\rho(a)x_\rho(a)) + (1-q_1)(p_\rho(b) - \mu_\rho(b)x_\rho(b)) \\
\text{s.t. } & \mu_\rho(b)x_\rho(b) - \frac{\eta}{2}(1-x_\rho(b))^2 - p_\rho(b) \geq -\frac{\eta}{2} + I_0 \\
& \mu_\rho(a)x_\rho(a) - \frac{\eta}{2}(1-x_\rho(a))^2 - p_\rho(a) \geq -\frac{\eta}{2} \\
& \mu_\rho(b)x_\rho(b) - \frac{\eta}{2}(1-x_\rho(b))^2 - p_\rho(b) \geq \mu_\rho(b)x_\rho(a) - \frac{\eta}{2}(1-x_\rho(a))^2 - p_\rho(a) \\
& \mu_\rho(a)x_\rho(a) - \frac{\eta}{2}(1-x_\rho(a))^2 - p_\rho(a) \geq \mu_\rho(a)x_\rho(b) - \frac{\eta}{2}(1-x_\rho(b))^2 - p_\rho(b)
\end{aligned}
$$

This problem is analogous to RS problem, with the exception that the RHS of type b's IR constraint has increased by I_0. This difference changes the set of restrictions that bind for each correlation.

Consider first a correlation $\rho < \rho^*$. In RS problem, type b's IR constraint and type's a IC constraint are the only that binds. This is still true in this case because only the RHS of type b's IR constraint has increased. Therefore, the optimal coverages are the same and the only difference is that now the insurer has to decrease both prices by I_0. Therefore, in this region the profit function is still increasing and convex.

If instead $\rho > \rho^*$, RS contract might no be feasible, because according to proposition 1 type b's utility in RS contract is equal to $u_\rho(b) = \mu_\rho(b) - \mu_\rho(a) - \frac{1-q_1}{\eta q_1}(\mu_\rho(b) - \mu_\rho(a))^2 - \frac{\eta}{2}$, and this problem requires that it is at least equal to I_0. Since u_ρ is strictly concave and by definition $u_{\mathbb{E}[\rho]} = I_0 - \frac{\eta}{2}$, there exist two thresholds $\tilde{\rho}_1 \leq \tilde{\rho}_2$ such that $u_\rho(b) < I_0 - \frac{\eta}{2}$ for $\rho \in (\rho^*, \tilde{\rho}_1) \bigcup (\tilde{\rho}_2, \bar{\rho})$ and $u_\rho(b) > I_0 - \frac{\eta}{2}$ for $\rho \in (\tilde{\rho}_1, \tilde{\rho}_2)$. Then type b's constraint IR should bind only for correlations $\rho \in (\rho^*, \tilde{\rho}_1) \bigcup (\tilde{\rho}_2, \bar{\rho})$.

Now suppose that none of the IC constraints bind. Then $x_a(\rho)$ and $x_\rho(b)$ would be both equal to 1 and $p_\rho(b) = \mu_\rho(b) + \frac{\eta}{2} - I_0$ and $p_\rho(a) = \mu_\rho(a) - \frac{\eta}{2}$. Comparing both prices allow us to find which IC constraint should bind: If $p_\rho(b) > p_\rho(a)$, ie, $\mu_\rho(b) - \mu_\rho(a) > I_0$, type b's IC constraint should bind and and in the opposite case type a's IC constraint should bind. We denote by $\hat{\rho}^*$ the unique correlation for which $\mu_{\hat{\rho}^*}(b) - \mu_{\hat{\rho}^*}(a) = I_0$. By definition of $u_\rho(b)$ and I_0 we conclude that $u_{\hat{\rho}^*} < I_0 - \frac{\eta}{2}$ and $\hat{\rho}^* < \mathbb{E}[\rho]$. Convexity of $u_\rho(b)$ and the fact that $u_{\mathbb{E}[\rho]} = I_0 - \frac{\eta}{2}$ imply that $\hat{\rho}^* < \tilde{\rho}_1$.

Therefore, for $\rho > \hat{\rho}^*$ we have three sets of correlations: $(\hat{\rho}^*, \tilde{\rho}_1)$, $(\tilde{\rho}_1, \tilde{\rho}_2)$ and $(\tilde{\rho}_2, \bar{\rho})$. In the first and third set, type b's IC and IR constraints and type a's IR constraint bind. In this case, the optimal coverages are $x_\rho(b) = 1$ and $x_\rho(a) = \frac{I_0}{\mu_\rho(b) - \mu_\rho(a)}$ and optimal prices are equal to $p_\rho(b) = \mu_\rho(b) + \frac{\eta}{2} - I_0$ and $p_\rho(a) = \frac{\mu_\rho(a)I_0}{\mu_\rho(b) - \mu_\rho(a)} + \frac{\eta}{2} - \frac{\eta}{2}\left(1 - \frac{I_0}{\mu_\rho(b) - \mu_\rho(a)}\right)^2$. Therefore, for these correlations profits are equal to

$$
\frac{\eta}{2} - (1-q_1)I_0 - q_1\frac{\eta}{2}\left(1 - \frac{I_0}{\mu_\rho(b) - \mu_\rho(a)}\right)^2.
$$

This function is clearly decresing since $I_0 < \mu_\rho(b) - \mu_\rho(a)$ and $\mu_\rho(b) - \mu_\rho(a)$ is increasing in ρ.

Further, this profit function is strictly convex iff $\mu_\rho(b) - \mu_\rho(a) > \frac{3}{2}I_0$, this is, this profit function is locally concave close to $\hat{\rho}^*$.

In the set $(\tilde{\rho}_1, \tilde{\rho}_2)$ only type b's IC constraint and type a's IR constraint bind. Then the optimal coverages and prices are exactly the same as in RS problem. Therefore, the profit function in this set is strictly decreasing and strictly convex.

Finally, we need to characterize the optimal contract for correlations in the set $(\rho^*, \hat{\rho}^*)$. If only type b's IR constraint and type a's IC constraint bind the optimal coverages would be $x_\rho(b) = 1 + \frac{q_1}{(1-q_1)\eta}(\mu_\rho(b) - \mu_\rho(a))$ and $x_\rho(a) = 1$ and prices would be $p_\rho(b) = \frac{\eta}{2} - I_0 + \mu_\rho(b)x_\rho(b) - \frac{\eta}{2}\left(1 - x_\rho(b)\right)^2$ and $p_\rho(a) = \frac{\eta}{2} - I_0 + \mu_\rho(a)x_\rho(a) + x_\rho(b)(\mu_\rho(b) - \mu_\rho(a))$. However, with this contract type a's utility is equal to $I_0 - (\mu_\rho(b) - \mu_\rho(a))x_\rho(b) - \frac{\eta}{2}$. Therefore, if $(\mu_\rho(b) - \mu_\rho(a)) + \frac{q_1}{(1-q_1)\eta}(\mu_\rho(b) - \mu_\rho(a))^2 > I_0$, which by definition is true close to ρ^*, constraint IR should bind. In that case $x_\rho(b) = \frac{I_0}{\mu_\rho(b) - \mu_\rho(a)}$, $p_\rho(b)$ is obtained from the same formula and $p_\rho(a) = \frac{\eta}{2} + \mu_\rho(a)$. Let $\tilde{\rho}_3$ be the unique correlation at which $(\mu_{\tilde{\rho}_3}(b) - \mu_{\tilde{\rho}_3}(a)) + \frac{q_1}{(1-q_1)\eta}(\mu_{\tilde{\rho}_3}(b) - \mu_{\tilde{\rho}_3}(a))^2 = I_0$.

Therefore, for $\rho \in (\rho^*, \tilde{\rho}_3)$ the structure of the solution is the same as the RS contract for correlations $\rho > \rho^*$. Then, in this set, the profit function is strictly increasing and convex. The profit function in the set $(\tilde{\rho}_3, \hat{\rho}^*)$ is equal to

$$\frac{\eta}{2} - (1 - q_1)I_0 - (1 - q_1)\frac{\eta}{2}\left(1 - \frac{I_0}{\mu_\rho(b) - \mu_\rho(a)}\right)^2,$$

which is strictly increasing since in this case $I_0 > \mu_\rho(b) - \mu_\rho(a)$, and as in the set $(\hat{\rho}^*, \tilde{\rho}_1)$ before, it is concave close to $\hat{\rho}^*$ and strictly convex far from it.

In summary, with competition the insurer's profit funtion has unique global and local maximum at the correlation $\hat{\rho}^*$, strictly increasing to the left of this correlation and strictly decreasing to its right. Further, it is locally concave close to that correlation, and strictly convex for the rest of the support.

Step 2. In this step, we use the argument in the proof of Proposition 2 to show that the insurer's optimal message function has the same structure that in the case without competition.

First, it is clear that the insurer wants to send the same message m for all correlations $\rho < \hat{\rho}^*$ and assigns for some correlations $\rho > \hat{\rho}^*$ in order to make sure that ones that insuree observes the message m, the insuree's posterior expected correlation is exactly equal to $\hat{\rho}^*$. As in the proof of Proposition 2 one can argue that to maximize the probability that is assign to this message $r(\rho) = m$ iff $\rho < \hat{\rho}^{**}$ with $\hat{\rho}^{**}$ the correlation defined as the unique correlation for wich $\mathbb{E}[\rho \mid \rho < \hat{\rho}^{**}] = \hat{\rho}^*$.

Now, by assumption $\mu_{\hat{\rho}^{**}}(b) - \mu_{\hat{\rho}^{**}}(a) > \frac{3}{2}I_0$. According to the first step, this condition guarantees that the profit function at any correlation $\rho > \hat{\rho}^{**}$ is strictly convex. Therefore the insurer wants to disclose all information for any correlation $\rho > \hat{\rho}^{**}$.

Therefore, we conclude that the first bullet-point in the statement is satisfied.

Step 3. Finally, we make some comparisons in terms of the insurer's profits, the insuree's information rent (or consumer surplus) and in terms of total surplus to see how efficiency is affected.

First, by strict convexity of the profit function without competition the optimal expected profits without competition for correlations $\rho < \hat{\rho}^{**}$ are larger than $\pi^{RS}(\hat{\rho}^{*})$. Since, with competition, at $\hat{\rho}^{*}$ both types IC constraints bind and type a's IR constrint bind, but withput competition, typoe a's IC constraint does not, the profits at this correlation with competition has to be strictly smaller than $\pi^{RS}(\hat{\rho}^{*})$. Further, for each correltion at which there is perfect information disclosure with competition, there is perfect disclosure withpout competition, and the increase of the outside option in type b's IR constraint can only reduce profits.

Second, total surplus is the same at the pooling regions because both with and without competition both types of insurees obtain a contract which provides full insurance. For correlations, $\rho > \hat{\rho}^{**} > \tilde{\rho}$, we show in Step 1, that with competition type b's coverage is equal to RS coverage in type's a coverage is either larger or equal than RS coverage, depending on whether type b's IR constraint binds or not. Therefore, in this region total surplus is weakly larger.

Third, consider a correlation ρ in the pooling region without competition. Type b's ex-post consumer surplus is equal to $\mu_r ho(b) - \mu_{\rho^*}(b)$ and type a's consumer surplus is equal to $\mu_r ho(a) - \mu_{\rho^*}(a)$, both without competition. When there is competition this correlation still belongs to the pooling region, type b's ex-post consumer surplus is equal to $\mu_r ho(b) - \mu_{\hat{\rho}^*}(b) + I_0$ and type a's consumer surplus is equal to $\mu_r ho(a) - \mu_{\hat{\rho}^*}(a)$. As $\hat{\rho}^* > \rho^*$, and $\mu_\rho(b)$ is increasing in ρ and $\mu_\rho(a)$ is decreasing in ρ, we inmediately conclude that type a's ex-post consumer surplus is always larger with compwtition and type b's ex-post consumer surplus is larger with competition iff $I_0 > \mu_{\hat{\rho}^*}(b) - \mu_{\rho^*}(b)$.

Finally, consider a correlation ρ in the full-disclosure region with competition, this is, $\rho > \hat{\rho}^{**} > \tilde{\rho}$. Type b's information rent is larger or equal with competition depending on whether type b's IR constraint binds or not. Type a's information rent is zero both and without competition. Therefore, with competition none of the types is worst off. □

References

L. Abrardi, L. Colombo, and P. Tedesch. The value of ignoring risk: Competition between better informed insurers. Politecnico di Torino andUniversit'a Cattolica del Sacro Cuore, 2022.

G. Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, 84(3):488–500, 1970.

D. J. Benjamin. Errors in probabilistic reasoning and judgment biases. In B. D. Bernheim, S. DellaVigna, and D. Laibson, editors, *Handbook of Behavioral Economics: Applications and Foundations 1*, volume 2, chapter 2, pages 69–186. Elsevier, 2019.

D. Bergemann and S. Morris. Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522, 2016.

J. Best and D. Quigley. Persuasion for the long run. forthcoming, *Journal of Political Economy*, 2023.

D. Bhaskar, A. McClellan, and E. Sadler. Regulation design in insurance markets. forthcoming, *Ameircan Economic Review*, 2023.

B. Biais, D. Martimort, and J. Rochet. Competing mechanisms in a common value environment. *Econometrica*, 68(4):799–837, 2000.

T. Börgers. *An Introduction to the Theory of Mechanism Design*. Oxford University Press, 2015.

Council of the European Union. Regulation 2016/679 of the European Parliament and of the Council: General Data Protection Regulation. eur-lex.europa.eu/eli/reg/2016/679/oj, *Official Journal of the European Union*, 2016.

V. P. Crawford and J. Sobel. Strategic information transmission. *Econometrica*, 50(6):1431–1451, 1982.

R. Eilat, K. Eliaz, and X. Mu. Bayesian privacy. Ben Gurion University and Tel Aviv University and Columbia University, 2020.

L. Einav and A. Finkelstein. Selection in insurance markets: Theory and empirics in pictures. *Journal of Economic Perspectives*, 25(1):115–138, 2011.

H. Fang and Z. Wu. Consumer vulnerability and behavioral biases. NBER, 2020.

H. Fang, M. P. Keane, and D. Silverman. Sources of advantageous selection: Evidence from the medigap insurance market. *Journal of Political Economy*, 116(2):303–350, 2008.

Financial Times. Data brokers: regulators try to rein in the 'privacy deathstars'. www.ft.com/content/f1590694-fe68-11e8-aebf-99e208d3e521, January 7, 2019.

A. Finkelstein and K. McGarry. Multiple dimensions of private information: Evidence from the long-term care insurance market. *American Economic Review*, 96(4):938–958, 2006.

G. R. Fréchette, A. Lizzeri, and J. Perego. Rules and commitment in communication: An experimental analysis. *Econometrica*, 90(5):2283–2318, 2022.

M. Gentzkow and E. Kamenica. A Rothschild-Stiglitz approach to Bayesian Persuasion. *American Economic Review P&P*, 106(5):597–601, 2016.

J. R. Green and J.-J. Laffont. *An Introduction to the Theory of Mechanism Design*. North-Holland Publishing Company, 1979.

S. J. Grossman and J. E. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70(3):393–408, 1980.

B. R. Handel and J. T. Kolstad. Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500, 2015.

E. Kamenica and M. Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

E. Kamenica, S. Mullainathan, and R. Thaler. Helping consumers know themselves. *American Economic Review P&P*, 101(3):417–422, 2011.

B. Keller, M. Eling, H. Schmeise, M. Christen, and M. Loi. Big data and insurance: Implications for innovation, competition and privacy. Technical report, International "Geneva" Association for the Study of Insurance Economics, www.genevaassociation.org/research-topics/cyber-and-innovation-digitalization/big-data-and-insurance-implications-innovation, 2018.

L. M. Khan. Amazon's antitrust paradox. *The Yale Law Jounral*, 126(3):710–805, 2017.

F. Koessler and V. Skreta. Selling with evidence. *Theoretical Economics*, 14(2):345–371, 2019.

J.-J. Laffont and D. Martimort. *The theory of incentives: the principal-agent model*. Princeton university press, 2009.

X. Lin and C. Liu. Credible persuasion. forthcoming, *Journal of Political Economy*, 2023.

V. F. Luz, P. Gottardi, and H. Moreira. Risk classification in insurance markets with risk and preference heterogeneity. forthcoming, *Review of Economic Studies*, 2023.

G. J. Mailath and A. Postlewaite. Asymmetric information bargaining problems with many agents. *Review of Economic Studies*, 57(3):351–367, 1990.

E. Maskin and J. Tirole. The principal-agent relationship with an informed principal: The case of private values. *Econometrica*, 58(2):379–409, 1990.

E. Maskin and J. Tirole. The principal-agent relationship with an informed principal ii: Common values. *Econometrica*, 60(1):1–42, 1992.

R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

R. B. Myerson. Optimal coordination mechanisms in generalized principal-agent problems. *Journal of Mathematical Economics*, 10(1):67–81, 1982.

R. B. Myerson. Mechanism design by an informed principal. *Econometrica*, 51(6):1767–1797, 1983.

T. Mylovanov and T. Tröger. Mechanism design by an informed principal: Private values with transferable utility. *Review of Economic Studies*, 81(4):1668–1707, 2014.

R. Rajan. Big technology companies are modern monopolies, but don't break them up. www.barrons.com/articles/big-tech-companies-modern-monopolies-google-facebook-51552076382, March 09, 2019.

E. Ramirez, J. Brill, M. K. Ohlhausen, J. D. Wright, and T. McSweeny. Data brokers: A call for transparency and accountability. Technical report, Federal Trade Commission, www.ftc.gov/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014, 2014.

J. Riley. Mathemtical note: Stochastic dominance. UCLA, 2017. URL http://www.econ.ucla.edu/riley/201C/2017/MathematicalNote.pdf.

M. Rothschild and J. Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics*, 90(4):629–649, 1976.

J. E. Stiglitz. Monopoly, non-linear pricing and imperfect information: The insurance market. *Review of Economic Studies*, 44(3):407–430, 1977.

R. Strausz. Correlation-savvy sellers. forthcoming, *RAND Journal of Economics*, 2023.

J. Tirole. Overcoming adverse selection: How public intervention can restore market functioning. *American Economic Review*, 102(1):29–59, 2012.

B. Villeneuve. Competition between insurers with superior information. *European Economic Review*, 49(2):321–340, 2005.