Inverse selection*

Markus K. Brunnermeier Rohit Lamba Carlos Segura-Rodriguez

February 2026

Abstract

Big data and AI invert classical adverse selection: insurers can now infer statistical correlations that consumers cannot, reversing the traditional informational advantage. We study insurance contracting where a two-dimensional state determines risk, the agent privately knows one dimension, and the insurer privately knows how the two correlate. The insurer faces a fundamental trade-off between obfuscation and price discrimination—fine-tuned contracts enable better screening but may reveal her statistical advantage. The optimal policy exhibits a bang-bang structure: pooling on a statistical model that renders the agent's information worthless, and full disclosure elsewhere. Profits increase substantially when agents fail to perform Bayesian inference.

1 Introduction

The rise of big data, artificial intelligence (AI), and machine learning is one of the defining characteristics of the 21st century economy. Companies record almost every characteristic of people, which can be correlated to better predict outcomes. However, most models in information economics assume that customers have an informational advantage over sellers. In these models, the principal, e.g., the insurance company, faces an adverse selection problem.[1]

Although customers may still have private information about some of their characteristics, big data allows insurance companies to develop superior aggregate information, using new statistical tools to better infer correlates about characteristics and ultimate risk. In other words, the principal here can "invert" the mapping from characteristics to risks. Thus, big data and AI transform many adverse selection problems into what we call the "inverse selection" problems.

*Brunnermeier: markus@princeton.edu, Princeton University; Lamba: rohitlamba@cornell.edu, Cornell University; Segura-Rodriguez: segurarc@bccr.fi.cr, Banco Central de Costa Rica. We are indebted to Dirk Bergemann (editor), Gabriel Chodorow-Reich (supporting editor), and five anonymous referees for their generous advice. We'd also like to thank Nemanja Antić, Levon Barseghyan, Tommaso Denti, Maryam Farboodi, 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]See Green and Laffont [1979] and Laffont and Martimort [2009] for general theoretical treatments of the principal agent screening problem.

To capture this, we use the informed principal approach within mechanism design (Myerson [1983] and Maskin and Tirole [1990, 1992]), but also depart from the canonical structure in two ways: First, the endowment of information to the principal and agent through different pieces of a two dimensional state, and second, the principal's commitment to an information disclosure policy and menu of contracts. Moreover, the basic structure here is inspired by 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.

The two dimensional type space that underlies the inverse selection problem determines the riskiness of the agent: The agent perfectly knows one characteristic, the first dimension of the type space. Think of this as hard private information—family history, eating habits, etc. The second dimension is a genetic condition whose exact realization is unknown to everyone in the model. The principal, e.g., the insurer, privately knows the correlation across both dimensions. Think of a statistical model that computes how these dimensions interact and jointly determine the agent's probability of, say, getting a disease.

The principal faces a fundamental tension: she must determine the optimal level of granularity in designing and fine-tuning the menu of contracts. Greater granularity enables more precise price discrimination; that is, distinct premium-coverage pairs for different statistical relations. However, offering highly tailored screening contracts may inadvertently disclose part of her statistical informational advantage, namely, the correlation between both characteristics. Thus, the principal confronts an inherent trade-off between obfuscation and price discrimination.[2]

The principal could, in principle, select a highly granular menu of contracts that fully reveals her information about the underlying correlation. Under such complete disclosure, the model effectively reduces to a monopolistic variant of the Rothschild and Stiglitz [1976] (RS) benchmark.[3] At the opposite extreme, the principal could entirely obfuscate her informational advantage by pooling and offering a single unconditional contract across all possible correlation structures. But such complete pooling would not allow the principal to appropriately exploit the informational advantage.

The optimal disclosure policy, contingent on the actual correlation, is characterized by an appropriate combination of these two extremes. Consider all the possible correlations on a number line and appropriately split it into two intervals. In one region, full disclosure—zero obfuscation—prevails, and the RS benchmark with price discrimination is implemented. In the complementary region, complete obfuscation through pooling eliminates price discrimination. Relative to the RS benchmark, the principal performs weakly better point-wise for all possible

[2] Note that this trade-off is different from the standard *rent-versus-efficiency* where the principal worsens efficient risk-sharing in order to minimize the information rent that the agent can extract. This rent-versus-efficiency trade-off is still present in our setting with respect to the agent's private information.

[3] The reduction to RS occurs with an important twist: the classification into "high" and "low" types depends jointly on the agent's private information and the statistical correlation observed by the principal.

statistical relations under inverse selection and strictly better in the pooling region by strategically choosing which statistical relationship to credibly communicate to the agent.

Inverse selection differs not only from the standard *adverse selection* but also from the more recent *advantageous selection* literature. Advantageous selection emphasizes preference heterogeneity in order to overturn the standard theoretical, but empirically counterfactual, result that high-risk agents get full insurance whereas low-risk agents opt for partial insurance. Now highly risk-averse agents buy more insurance, despite the fact that they are less risky.[4] In both settings, adverse and advantageous selection, the insurance provider suffers from an informational disadvantage. This contrasts with inverse selection, which may thus be regarded as a third generation of models.

The formal analysis is based on the information design (Kamenica and Gentzkow [2011] and Bergemann and Morris [2016]) and contract design (as in Börgers [2015], Chapter 2) problems, wherein the principal can fully specify at the outset the menu of contracts and what information they disclose to the agent.[5] It also assumes that the agent is fully rational and capable of Bayesian inference, which imposes substantial constraints on the principal's profits. Yet such an assumption is behaviorally unrealistic in many real-world settings (see Barseghyan, Molinari, O'Donoghue, and Teitelbaum [2013] and Handel and Kolstad [2015]). In reality, agents may lack not only probabilistic sophistication but also the cognitive ability to reconstruct the mapping from the principal's contract offering to payoff-relevant beliefs.

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 how the insurer could create maximal obfuscation and implement maximal price discrimination if she could simply shape the belief of the insuree. Second, we study the case in which the insuree is *naive*, i.e., he again does not infer any statistical information from the contract, but unlike the gullible, sticks to the prior. For the naive case, the difference between the insurer's and insuree's beliefs is exogenously fixed by the prior, and the insurer maximizes the price discrimination channel.

Finally, in the supplementary appendix, we explore two extensions: First, we introduce competition into the model by allowing other "regular" insurers, who do not have access to the big data technology, to offer contracts. They offer a single Rothschild-Stiglitz contract, averaging over all possible correlations, in order to screen the insuree along the first dimension. We show

[4] Einav and Finkelstein [2011] provide an overview of the key ideas, and Finkelstein and McGarry [2006] and Fang, Keane, and Silverman [2008] empirical evidence.

[5] The information structure here also relates to mechanism design with correlated information [Crémer and McLean, 1988, Riordan and Sappington, 1988], where correlation helps an uninformed principal cross-check the agent's report and reduce information rents. The key difference here is that the principal *privately* knows the correlation structure, transforming the problem into one where she strategically manages disclosure—introducing the obfuscation-versus-discrimination tradeoff central to our analysis.

that competition reduces some of the informational advantage of the insurer and increases both the surplus of the agent and total welfare. Second, we examine the setting in which the insurer is forced to put the statistical information (its big data advantage) in the public domain. This too increases the insuree's surplus; however, contrary to competition, it lowers the insurer's profit and total surplus.

While our model is admittedly stylized, it provides a framework for policymakers 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 has manifested in the rise of data brokers.[6] Second, forcing the principals to make private statistical information public points towards the merits of a public data repository.[7] On the other hand, making consumers 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.[8] Finally, competition is another way to protect the consumer as it reduces the extent of price discrimination by big data monopolies.[9]

To our knowledge, Villeneuve [2005] is the first article to study insurance markets in the realm of the informed principal model. This has been followed up by Abrardi, Colombo, and Tedesch [2022], simultaneously, with our work. Both of these papers, however, 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.

In recent work, Luz, Gottardi, and Moreira [2023] and Bhaskar, McClellan, and Sadler [2023] also examine a two-dimensional type space for insurance contracts. The former considers heterogeneity in preferences, specifically risk aversion, as the second dimension, while the latter assumes that the first dimension is commonly known and can be used by a third party, such as a regulator, to implement the efficient outcome. In price discrimination settings, Strausz [2023] studies how a multi-product seller can discriminate based on buyers' valuations by constructing correlations using big data, and Deb, Pai, and Roesler [2024] studies how informed the buyer would want the seller to be about the underlying value from bilateral trade. Also related is Kamenica, Mullainathan, and Thaler [2011], which explores the consequences of a firm knowing more about the value of a product than the consumers themselves.

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

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

[8] 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.

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

2 Model

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 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 risk averse and is assumed to have a standard mean-variance preference:

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

where $\mu = \mathbb{E}[\ell]$ is the expected loss, $\mathbb{V}[\ell]$ the variance of loss, γ measures the extent of risk aversion, and $\eta = \gamma \times \mathbb{V}[\ell]$ captures the level of risk faced by the insuree. This lends a tractable structure to the insuree's utility—it is linear in money and concave in the extent of loss.[10] The insurer has a standard linear preference over price and expected payment:

$$\pi(p, x) = p - \mu x.$$

Information. We consider a two-dimensional state $\boldsymbol{\theta} = (\theta_1, \theta_2)$, where each component $\theta_i \in \{a, b\}$ is binary. Interpret θ_1 as the agent's lifestyle type—disciplined habits versus indulgent ones—something he knows but that is difficult for the insurer to verify. The second dimension, θ_2, captures latent physiological vulnerability (e.g., inflammatory predisposition) that the agent does not directly observe. The insurer, however, can statistically infer this from population-level patterns in purchasing behavior, healthcare utilization, or wearable data.

The outcome μ reflects health risk as expected loss; both dimensions contribute, with the combination of indulgent lifestyle and high vulnerability generating the highest risk. The principal's big data advantage lies in knowing how these dimensions interact: the correlation ρ captures whether unhealthy lifestyles co-occur with physiological vulnerability, information the agent cannot estimate from his own experience.

Formally, $\mu_{\boldsymbol{\theta}}$ can assume up to four distinct values depending on the realizations of θ_1 and θ_2, sorted as follows:

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

The joint distribution of θ is given by $q = (q_{aa}, q_{ab}, q_{ba}, q_{bb})$, and let $q_1 = q_{aa} + q_{ab}$ and $q_2 = q_{aa} + q_{ba}$ be the marginal distributions of θ_1 and θ_2, respectively. Let ρ be the correlation

[10] A standard behavioral foundation in related models for the mean-variance preference is the CARA-Gaussian framework; see, for example, the seminal Grossman and Stiglitz [1980].

coefficient between θ_1 and θ_2, and define $\sigma = \sqrt{q_1(1-q_1)}\sqrt{q_2(1-q_2)}$. As shown below, the joint distribution can then be rewritten using three parameters: ρ, q_1, q_2.

$$\begin{array}{cc|c|c|c} & & & \theta_2 & \\ & & a & b & \\ \hline \theta_1 & a & q_1 q_2 + \rho\sigma & q_1(1-q_2) - \rho\sigma & q_1 \\ \hline & b & (1-q_1)q_2 - \rho\sigma & (1-q_1)(1-q_2) + \rho\sigma & 1-q_1 \\ \hline & & q_2 & 1-q_2 & \end{array}$$

The insuree privately observes θ_1 and knows the marginal distribution of θ_2, and the insurer exclusively knows the joint distribution of θ—this non-standard endowment of information in an insurance problem is what leads to the inverse selection structure. The marginals q_1 and q_2 are common knowledge, and to close the model, we assume that ρ is drawn from F on $[\underline{\rho}, \overline{\rho}]$, where F has a continuous density f, and is commonly known.[11]

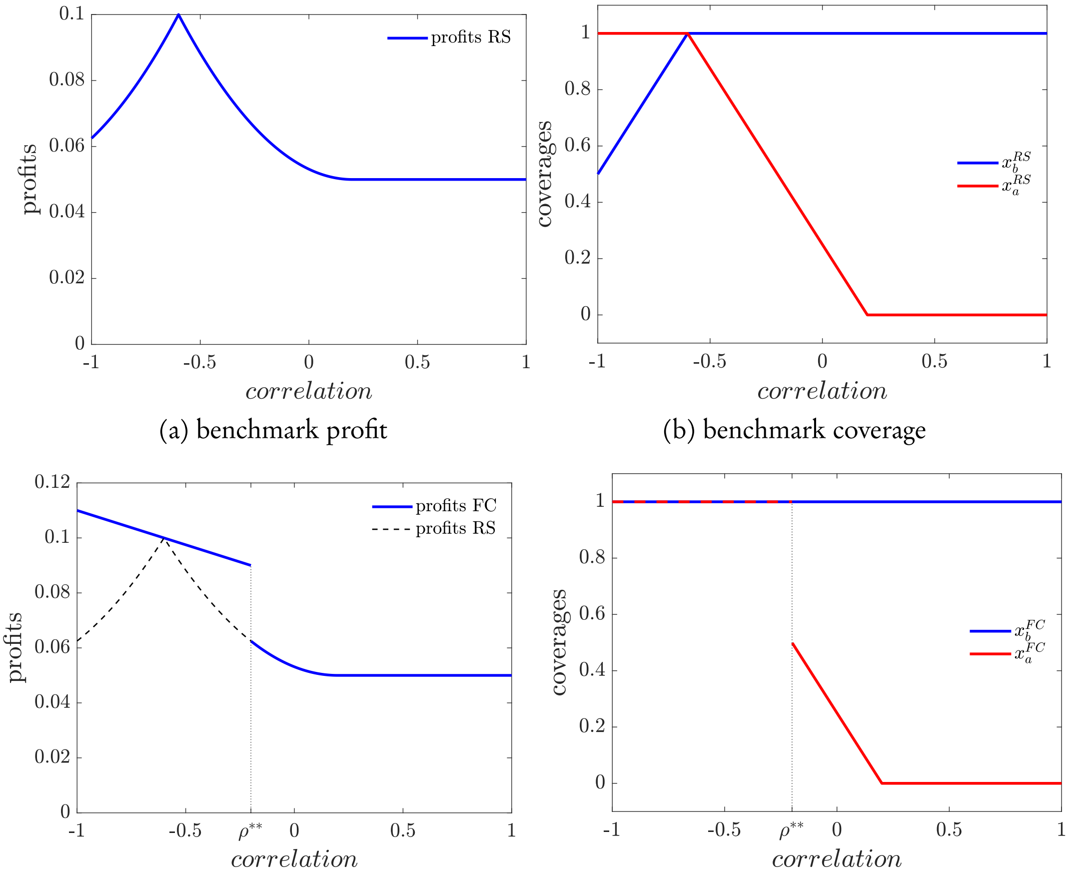
In the terminology of regression analysis, consider a principal (the insurer) who is capable of perfectly estimating the regression coefficients associated with θ_1 and θ_2 based on historical patient data. Nevertheless, the principal lacks knowledge of the specific realizations of θ_1 and θ_2 for the new agent under consideration. By contrast, the rational agent (insuree) possesses knowledge of θ_1 and recognizes that his estimates are subject to omitted variable bias. Consequently, he seeks to infer the parameter ρ from the principal to mitigate this bias.

Timing. Nature draws ρ from F on $[\underline{\rho}, \overline{\rho}]$, then θ_1 and θ_2 from the joint distribution defined by (q_1, q_2, ρ). The principal commits to a menu of contracts C, and correspondingly a disclosure policy about ρ. The agent observes θ_1, updates beliefs about ρ from the principal's offered contract, and then accepts or rejects it. Uncertainty is resolved, determining μ, and payoffs are realized.

The formal question. The question we ask is: What is the principal's optimal contract? Two observations: this is a standard price discrimination problem in which the principal posts a menu for the agent to self-select, and the menu may reveal information about ρ before selection. Thus, the problem becomes one of joint information and contract design.

Assumption. Let $\mu_\rho(\theta_1)$ be the expected value of μ based on the realized value of ρ and θ_1, and let ρ^* solve the equation $\mu_\rho(b) = \mu_\rho(a)$, which is linear in ρ. We assume that $\rho^* \in (\underline{\rho}, \overline{\rho})$. This makes the insurer's profit in the canonical Rothschild-Stiglitz version of our problem single-peaked in ρ (see Section 3).

[11]Once the marginals are fixed to be q_1 and q_2, the set of feasible correlations is $[\underline{\rho}, \overline{\rho}] \subset [-1, 1]$, where $\overline{\rho} = \min\left\{\frac{q_1(1-q_2)}{\sigma}, \frac{q_2(1-q_1)}{\sigma}\right\}$ and $\underline{\rho} = \max\left\{-\frac{q_1q_2}{\sigma}, -\frac{(1-q_1)(1-q_2)}{\sigma}\right\}$.



(a) benchmark profit (b) benchmark coverage

(c) optimal profit (d) optimal coverage

Figure 1: Benchmark (Section 3) and Optimal (Section 4) contracts

3 Benchmark: No inverse selection

We start the analysis by looking at the benchmark model where ρ, the correlation coefficient, is common knowledge. So, the principal has no statistical informational advantage; hence, there is no scope for inverse selection. The problem becomes isomorphic to the monopolistic version of the classical Rothschild and Stiglitz [1976] problem, studied first by Stiglitz [1977]. Then, for any fixed value of ρ, the principal offers a menu of contracts:

$$C(\rho) = \{c_\rho(a), c_\rho(b)\}, \text{ where } c_\rho(\theta_1) = (p_\rho(\theta_1), x_\rho(\theta_1)) \text{ for } \theta_1 = a, b.$$

to separate the "high" risk agents from the low "risk" ones. Note, however, 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. Both parties take expectations over θ_2 using a joint distribution, and the insuree is incentivized to reveal θ_1 truthfully while ensuring that he wants to accept the contract.[12] We will refer to this as the *benchmark model* and label it RS, pointing to the classical

[12]The formal optimization problem is defined in Section 4.1 later; since the problem becomes routine when ρ is common knowledge, we directly present the main result.

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 solve $\mu_\rho(b) = \mu_\rho(a)$.[13] 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 1b). 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 1a).[14]

Naturally, the profit function is also single-peaked at ρ^* for in either direction the private information of the insuree becomes (statistically) more informative as ρ gets further away from ρ^*. The two pieces are convex; i.e., the drop in profit occurs at a decreasing rate.

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

$$x_\rho(b) = 1 \text{ and } x_\rho(a) = 1 - \frac{1-q_1}{\eta q_1}\Delta < 1^{15} \tag{1}$$

and the optimal profit is given by

$$\pi^{RS} = \frac{\eta}{2} - (1-q_1)\Delta + \frac{(1-q_1)^2}{2\eta q_1}\Delta^2, \tag{2}$$

where $\Delta = (\mu_\rho(b) - \mu_\rho(a))$. The first term of π^{RS} is the efficient surplus from full insurance; the second term is the information rent paid to the high type—linear and negative in Δ; the third term is the gain from screening—quadratic and positive in Δ. So as ρ moves away from ρ^*, increasing Δ, the gains from screening are quadratic, partially offsetting the linear growth in information rent, which explains the convex structure.

The novel twist here is that the notion of a high or low risk type is endogenous to the realization of ρ. This already indicates what may happen when ρ is privately observed by the insurer. She has two potential targets: first, to make the agent believe that $\rho = \rho^*$, where his private information becomes uninformative and thus worthless; second, to maximally invert the agent's

[13] Note from the Assumption in Section 2, ρ^* is interior, that is, $\rho^* \in (\underline{\rho}, \overline{\rho})$.

[14] 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.

[15] For a large enough distortion, the "low" type is shut out of the market, as seen in Figure 1b for high values of ρ.

understanding of risk—when ρ is large, making the insuree believe ρ is small, and vice versa. The first target dominates in the rational model, where Bayesian inference constrains feasible beliefs; the second dominates in the behavioral gullible model, where no such constraint binds.

4 Optimal information and contract design

Here, we study the joint information and contract design problem wherein the insurer chooses the information disclosure policy for ρ 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.

4.1 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 $r : [\underline{\rho}, \overline{\rho}] \to \Delta(\mathcal{M})$, where $\mathcal{M}$ is an arbitrary message space chosen as part of contract design, represents how coarsely (or finely) the insurer wants to communicate her information about the correlation coefficient to the insuree. For example, if r assumes only one value, no information beyond the prior is communicated, or if r assumes strictly monotone values, then information is fully disclosed, and so on.

Furthermore, 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:

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

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 rationality for the insuree.

The goal now is to characterize the optimal choice of (r, C). To that end, we now carefully define the objective and constraints of the optimization problem. Let $\pi(\rho)$ be the (ex post) profit of the insurer if her type is ρ. Then the (ex ante) objective 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:

$$\begin{aligned} u_m(\hat{\theta}_1; \theta_1) &= w - p_m(\hat{\theta}_1) - \left[1 - x_m(\hat{\theta}_1)\right] \mu_m(\theta_1) - \frac{\eta}{2}\left[1 - x_m(\hat{\theta}_1)\right]^2 \\ &= \underbrace{w - \mu_m(\theta_1)}_{\text{function of } \theta_1} + \underbrace{\left[x_m(\hat{\theta}_1)\mu_m(\theta_1) - \frac{\eta}{2}\left\{1 - x_m(\hat{\theta}_1)\right\}^2\right]}_{\text{function of } \theta_1 \hat{\theta}_1} - p_m(\hat{\theta}_1) \end{aligned} \tag{3}$$

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 a fixed message m drawn from $r(\rho)$, the insurer's profit is:

$$\pi(\rho) = q_1\left[p_m(a) - \mu_\rho(a)x_m(a)\right] + (1 - q_1)\left[p_m(b) - \mu_\rho(b)x_m(b)\right] \tag{4}$$

where $\mu_\rho(\theta_1)$ is the expected value of μ based on the realized value of ρ and the (truthfully) reported value of θ_1.

Two types of constraints are imposed on the optimization problem. First is the incentive constraint for the insuree (agent), that he 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 above, the insuree's incentive constraint incorporates the message coming from the mediator through the mapping r, by conditioning the (expected) utility on 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 (3). The information plus mechanism design problem then reads as follows:

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

4.2 The optimal information structure and contract

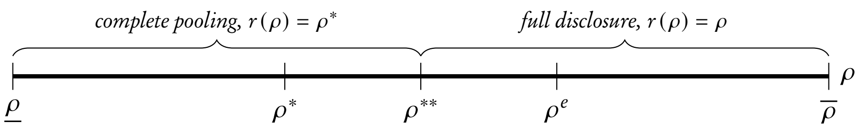
Following the burgeoned literature on Bayesian persuasion and information design (Kamenica and Gentzkow [2011] and Bergemann and Morris [2016]), the principal's (or the insurer's) problem can be considered as choosing posterior beliefs that satisfy the martingale (or Bayes' plausibility) condition, and then determining which contract to offer at each of those posterior beliefs. Since both the insurer and insuree's payoffs are linear in ρ, 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 as one in which m, and hence $r(\rho)$, coincides with the posterior mean $\hat{\rho}$ it induces.

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 1a, this function is convex on both sides of ρ^* and has a global maximum at ρ^*.

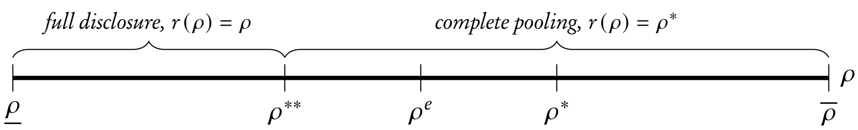
So the solution to problem $(\mathcal{P})$ boils down to which amongst the set of posterior beliefs that satisfy Bayes'rule will maximize the seller's expected profits, given that the RS-contract will be implemented at each of those (expected) beliefs in the second contract design stage. The solution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

Proposition 2. *Suppose $\rho^* \in (\underline{\rho}, \overline{\rho})$ 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 the 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 upon the insurer to satisfy the Bayes' plausibility condition, it is optimal to disclose full information due to the convexity of the profit function.

Put differently, 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 conduct Bayesian inference correctly, and in the other region, she discloses it all to achieve 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. Figures 1c and 1d depict the pointwise profit and coverage in the optimal contract as a function of ρ. In the pooling region, the RS contract corresponding to correlation ρ^* is offered, which provides full coverage. Since the insuree's expected correlation is ρ^*, he has no informational advantage whatsoever; hence, the full surplus is extracted. The profit ranking is summarized in the following simple corollary:

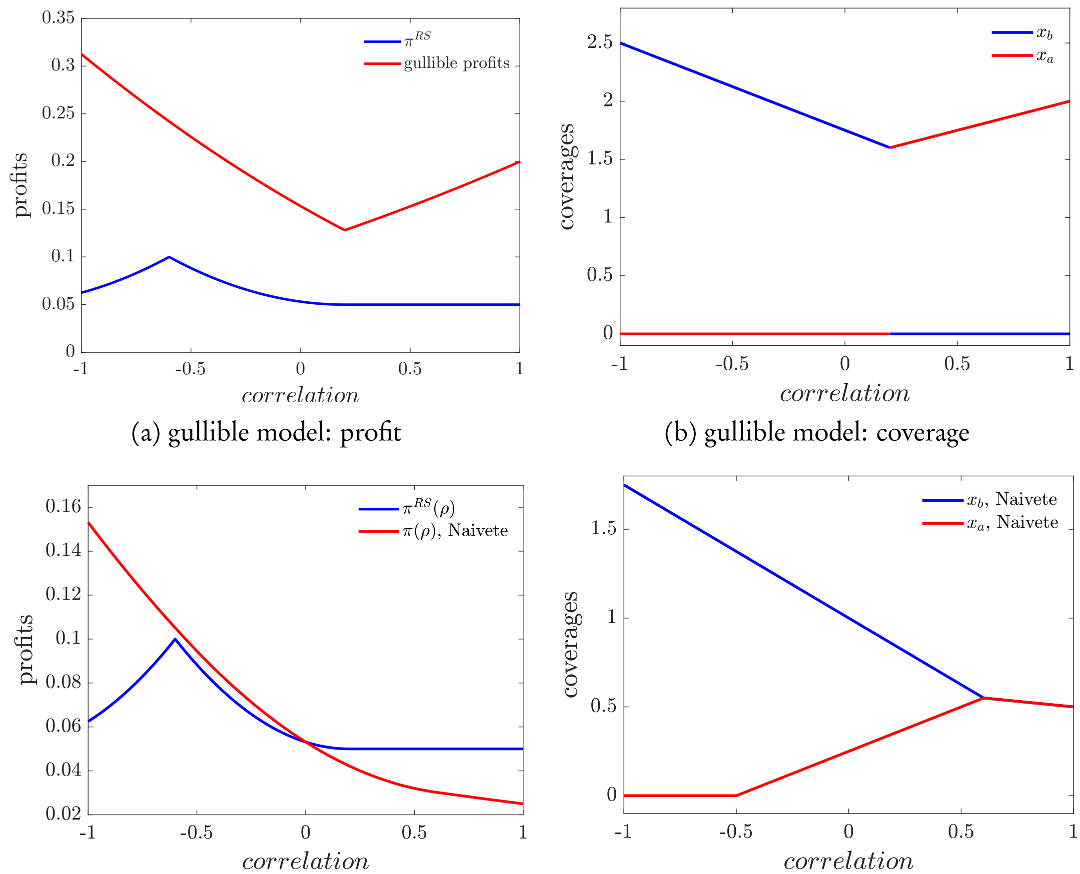
Corollary 1. *Let $\pi^{IS}(\rho)$ denote the insurer's profit under optimal inverse selection and $\pi^{RS}(\rho)$ the benchmark profit. Then $\pi^{IS}(\rho) \geq \pi^{RS}(\rho)$ for all ρ, with strict inequality in the pooling region for $\rho \neq \rho^*$.*

5 Behavioral approach: incomplete inference from contracts

The requirement of Bayesian inference 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 (see Benjamin [2019]). In 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, Barseghyan, Molinari, O'Donoghue, and Teitelbaum [2013], Handel and Kolstad [2015], and others argue that standard insurance models miss this difficulty. We explore two departures from perfect Bayesian inference; formal statements and proofs are in Brunnermeier, Lamba, and Segura-Rodriguez [2026b].

Gullible insuree. As a first case, suppose the insuree is gullible to the extent that he believes the correlation ρ told to him by the insurer. This stylized model 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. So, the insurer chooses r and C in tandem to create both maximal obfuscation and maximal price discrimination.

The optimal contract features binary messages: the insurer reports the extreme negative correlation, $\underline{\rho}$, when the actual correlation is high, and the extreme positive correlation, $\overline{\rho}$, when it is low. This binary communication pivots around a threshold, say $\tilde{\rho}$. When $\rho > \tilde{\rho}$, so that type $\theta_1 = b$ is likely to suffer a larger loss, the insurer reports the opposite and overinsures $\theta_1 = a$ while underinsuring $\theta_1 = b$. The reverse occurs when $\rho < \tilde{\rho}$. 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 to the type who actually has a high probability of loss. Figures 2a and 2b depict the profit and coverages; notably, profits are uniformly higher than the benchmark for all ρ.



(a) gullible model: profit

(b) gullible model: coverage

(c) naive model: profit

(d) naive model: coverage

Figure 2: Behavioral inverse selection

Naive insuree. Another standard behavioral assumption is that the agent ignores signals offered by the contract about the correlation coefficient, so that the posterior always equals the prior. In this situation, the role of r is moot: the insurer designs the contract as a function of ρ knowing the insuree will evaluate payoffs using the prior F.

In the naive model, obfuscation is determined exogenously by the gap between the fixed prior and realized ρ, and the insurer cannot influence it. This sometimes works in her favor, sometimes against. Suppose F implies the insuree believes $\theta_1 = b$ is the high-risk type. When the realized correlation is close to $\underline{\rho}$, the insurer profits by selling substantial coverage to $\theta_1 = b$, who is actually low-risk but believes otherwise (see left part of Figure 2d). When the realized correlation is close to $\overline{\rho}$, however, the insurer would like to sell generously to $\theta_1 = a$, but the incentive constraint $x_\rho(b) \geqslant x_\rho(a)$ binds; nor can she sell much to $\theta_1 = b$, who underestimates his risk. Thus, for extreme correlations, pooling is forced. Overall, the naive insurer earns higher profits on average than an uninformed seller offering the RS contract at $\mathbb{E}(\rho)$, but unlike the gullible case, this ranking is not uniform (see Figure 2c).

Dissecting the key forces. The coverages vary as a function of ρ (the insurer's private information) and θ_1 (the insuree's private information). The latter generates the classical rent-versus-efficiency tradeoff; the former generates the novel tension between obfuscation and price discrimination. In the benchmark, obfuscation is zero and price discrimination is exogenously determined by ρ. In the gullible case, both are endogenously chosen to maximize profit—there is no tradeoff. In the naive case, obfuscation is exogenous while price discrimination is endogenous, helping the insurer on average but not uniformly. In the rational model, Bayesian inference constrains feasible beliefs, washing away some gains from inverse selection, yet the insurer still extracts higher profits than the benchmark by pooling around ρ^*.

6 Final remarks

Motivated by big data and AI, this paper introduces a private statistical informational advantage for the insurer in an otherwise standard screening problem. By inferring correlations using sophisticated statistical models, the insurer can significantly increase profits, especially when insurees cannot perform Bayesian inference.

While our baseline model employs mean-variance preferences and a 2×2 type space, the mechanism appears qualitatively robust. The key structural feature is that profit is single-peaked in the posterior expectation of the correlation, attaining a maximum where the agent's information is uninformative and convex on either side. This creates incentives to pool mass at the peak and fully disclose elsewhere. The logic extends to environments with multiple local maxima—pooling at beliefs where buyer heterogeneity becomes uninformative, separation elsewhere.[16]

[16]Numerical simulations with CRRA preferences confirm the profit function retains this shape. For general type

A natural next step is to limit the commitment to information disclosure that the principal has here. The most obvious conceptual way to do that is to introduce an incentive constraint for the insurer as well in reporting her type to the mediator. This introduces significant technical challenges—analogous to why cheap talk models (in the spirit of Crawford and Sobel [1982]) are harder to solve than models of information design. In ongoing work, Brunnermeier, Lamba, and Segura-Rodriguez [2026a], we find that a lack of commitment significantly curtails the ability of the principal to benefit from the extra statistical information, but only if the agent can do perfect Bayesian inference.

Testable implications emerge from the model's structure. First, if insurers possess statistical advantages and insurees are rational, contract menus should exhibit clustering consistent with partial pooling rather than full separation. Second, when consumers fail to infer information from contracts, coverage should be negatively correlated with true risk: generous coverage sold to those least likely to claim. Third, competition among insurers with heterogeneous statistical capabilities should increase total coverage, while transparency mandates need not. Finally, rejection of the rational model would manifest in consumers failing to update beliefs in response to menu variations that reveal the insurer's information.

The ideas developed here can be applied to contexts other than insurance. For example, in credit markets, new fintech allows the credit issuing agency to obtain statistical information about the creditworthiness of a client, separate from the client being aware of hard information regarding their financial circumstances. Finally, greater statistical information on the side of the principal may encourage more market concentration; endogenizing data collection and market size is a promising question for future work.

7 Appendix

We present here the proofs of Propositions 1 and 2. The analysis of propositions for behavioral extensions can be found in the Supplementary Appendix, Brunnermeier, Lamba, and Segura-Rodriguez [2026b].

7.1 Proof for benchmark model in Section 3

Proof of Proposition 1. For a specific (commonly known) ρ the insurer's profit is given by

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

Her problem is to choose a contract $c_\rho = \{p_\rho(\theta_1), x_\rho(\theta_1)\}_{\theta_1=a,b}$ to maximize $\pi(\rho)$ subject to the feasibility constraints:

spaces $\Theta_1 \times \Theta_2$, profit depends on the joint distribution only through the conditional expectations $\mathbb{E}[\mu(\theta) \mid \theta_1]$. Profit is maximized where θ_1 is conditionally uninformative about expected loss; the optimal policy pools at such beliefs, though the geometry may be richer than simple intervals.

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

For $\rho = \rho^*$, the correlation for which $\mu_\rho(b) = \mu_\rho(a)$, there is no asymmetric information. Therefore, she maximizes profits by offering a pooling full insurance contract $x_{\rho^*}(b) = x_{\rho^*}(a) = 1$ and binds the IR constraints to extract expected surplus.

Suppose that $\rho > \rho^*$, that is, $\mu_\rho(b) > \mu_\rho(a)$. 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 Laffont and Martimort [2009]) to show that in this case IC_H and IR_L bind, and IC_L and IR_H are slack. Setting up the appropriate Lagrangian, taking the first order conditions, and simplifying, we get:

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

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

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

7.2 Proof for Section 4

Proof of Proposition 2. Here, we use $f(\rho, m)$ to denote the joint distribution over primitive types ρ and messages m generated by the information structure, and $f(m|\rho)$ for the corresponding conditional distribution.

Step 1. First, suppose the insurer sends messages with non-zero measure that generate posterior beliefs $\rho' < \rho^* < \rho''$. By the definition of ρ^*, the insurer's profit satisfies $\pi^{RS}(\rho^*) > \pi^{RS}(\rho')$ and $\pi^{RS}(\rho^*) > \pi^{RS}(\rho'')$. The insurer can thus increase profits by reassigning mass from the messages generating ρ' and ρ'' to a new message $\hat{m}$ that induces the posterior ρ^*. Since $\rho' < \rho^* < \rho''$, this reassignment can be constructed to be Bayes-plausible. This process can be repeated until the measure of messages generating posteriors on one side of ρ^* (either for correlations below ρ^* or above ρ^*) is zero. Therefore, **there cannot be a non-zero mass of posteriors both below and above ρ^*.**

Second, we show that **the function π^{RS} is strictly convex** on either side of ρ^* where the Rothschild-Stiglitz (RS) solution is interior, and constant otherwise. Suppose $\rho > \rho^*$, implying $\Delta = \mu_\rho(b) - \mu_\rho(a) > 0$. Substituting the RS solution into the insurer's profit function yields Equation (2) from Section 3: $\pi^{RS} = \frac{\eta}{2} - (1 - q_1)\Delta + \frac{(1-q_1)^2}{2\eta q_1}\Delta^2$. This function is strictly convex in Δ, which itself is increasing in ρ. Conversely, at a corner solution, the coverage offered is independent of ρ, and the insurer appropriates the entire (constant) surplus. Notice that the solution is interior in a right-side neighborhood of ρ^*.

Step 2. Consider the trivial case where $\rho^* = \mathbb{E}[\rho]$. The insurer can generate the posterior belief ρ^* with a single, uninformative message. Since π^{RS} is maximized at ρ^*, this strategy is optimal.

Step 3. Suppose $\mathbb{E}[\rho] > \rho^*$. Step 1 implies the information structure can only generate posteriors at ρ^* and to one side of it. Since $\mathbb{E}[\rho] > \rho^*$, Bayes plausibility requires that the support of the posterior beliefs must be larger than or equal to ρ^*.

Suppose there exists a set of messages $M \subset \mathcal{M}$ such that for any $m \in M$, $\int_{\underline{\rho}}^{\rho^*} f(m \mid \rho')f(\rho')d\rho' > 0$ and $\mathbb{E}[\rho \mid m] > \rho^*$. This means message m (which induces a posterior larger than ρ^*) is sent with positive probability by some types $\rho < \rho^*$.

Take any such message $m' \in M$. Bayes plausibility allows m' to be split into two new messages, $\hat{m}$ and $\tilde{m}$, inducing posteriors ρ^* and $\tilde{\rho} > \mathbb{E}[\rho \mid m']$, respectively. From Step 1, π^{RS} is strictly convex in a neighborhood of ρ^*. By Jensen's inequality, this new information structure improves the insurer's profit. This process can be repeated until no message m with $\mathbb{E}[\rho \mid m] > \rho^*$ is sent with positive probability by any type $\rho \leq \rho^*$.

Therefore, **with probability one, any type $\rho \leqslant \rho^*$ must send a message that generates the posterior belief ρ^*.**

Step 4a. Define ρ^{**} such that $\rho^* = \mathbb{E}[\rho \mid \rho \leq \rho^{**}]$. We claim that for $\rho \in (\rho^*, \rho^{**}]$, the insurer also sends, with probability one, a message generating the posterior ρ^*.

Suppose, for contradiction, there is a set of messages M such that for any $m \in M$, $\int_{\rho^*}^{\rho^{**}} f(m \mid \rho)f(\rho)d\rho > 0$ and $\mathbb{E}[\rho \mid m] > \rho^*$. This means some types in $(\rho^*, \rho^{**}]$ are not pooled to ρ^*.

From Step 3, all types $\rho \leqslant \rho^*$ are pooled to ρ^*. By the definition of ρ^{**}, to maintain Bayes plausibility, there must be another set of messages M' such that for $m' \in M'$, $\mathbb{E}[\rho \mid m'] = \rho^*$ and this message is sent by types above ρ^{**} (i.e., $\int_{\rho^{**}}^{\bar{\rho}} f(m' \mid \rho')f(\rho')d\rho' > 0$). This construction is necessary to balance the types in M that are not pooled to ρ^*.

Step 4b. We propose a new information structure to improve insurer profits. Take a message $m' \in M'$ and move an ϵ proportion of its density $f(\rho', m')$ for each $\rho' \in [\rho^{**}, \bar{\rho}]$ to a message $m \in M$. Simultaneously, move a δ proportion of the density $f(\rho, m)$ for each $\rho \in [\rho^*, \rho^{**}]$ from m to m'. We choose ϵ small enough to ensure $\delta < 1$, with $\delta = \epsilon \frac{E(\rho',m') - \rho^* G(\rho',m')}{E(\rho,m) - \rho^* G(\rho,m)}$, where: $E(\rho', m') = \int_{\rho^{**}}^{\bar{\rho}} \rho' f(\rho', m')d\rho'$, $G(\rho', m') = \int_{\rho^{**}}^{\bar{\rho}} f(\rho', m')d\rho'$, $E(\rho, m) = \int_{\rho^*}^{\rho^{**}} \rho f(\rho, m)d\rho$, and $G(\rho, m) = \int_{\rho^*}^{\rho^{**}} f(\rho, m)d\rho$

It is immediate that $\frac{E(\rho',m')}{G(\rho',m')} > \frac{E(\rho,m)}{G(\rho,m)} > \rho^*$ (these ratios are the conditional expectations). By definition of δ, it follows that $\delta G(\rho, m)\left(\frac{E(\rho,m)}{G(\rho,m)} - \rho^*\right) = \epsilon G(\rho', m')\left(\frac{E(\rho',m')}{G(\rho',m')} - \rho^*\right)$. This implies $\delta G(\rho, m) > \epsilon G(\rho', m')$, meaning more probability mass is moved to m' than to m.

Next, we prove this reassignment preserves $\mathbb{E}[\rho \mid m'] = \rho^*$ and increases $\mathbb{E}[\rho \mid m]$. First, let $A' = \int_{\underline{\rho}}^{\bar{\rho}} \rho' f(\rho', m')d\rho' - \epsilon E(\rho', m')$ and $B' = \int_{\underline{\rho}}^{\bar{\rho}} f(\rho', m')d\rho' - \epsilon G(\rho', m')$, the expected correlation and total mass of message m' after removing an ϵ portion from this message to message m. By definition:

$$\rho^* = \mathbb{E}[\rho \mid m'] = \frac{A' + \epsilon E(\rho', m')}{B' + \epsilon G(\rho', m')}$$
$$\Rightarrow B'\rho^* = A' + \epsilon \left(E(\rho', m') - \rho^* G(\rho', m')\right)$$
$$= A' + \delta \left(E(\rho, m) - \rho^* G(\rho, m)\right)$$

This implies $\rho^* = \frac{A' + \delta E(\rho,m)}{B' + \delta G(\rho,m)}$. Thus, the posterior for m' remains ρ^* after the reassignment.

Second, analogously to A' and B', let $A = \int_{\underline{\rho}}^{\bar{\rho}} \rho f(\rho, m) d\rho - \delta E(\rho, m)$ and $B = \int_{\underline{\rho}}^{\bar{\rho}} f(\rho, m) d\rho - \delta G(\rho, m)$. The original posterior is $\tilde{\rho} = \mathbb{E}[\rho \mid m] = \frac{A + \delta E(\rho,m)}{B + \delta G(\rho,m)}$. This implies that

$$B\tilde{\rho} = A + \delta \left(E(\rho, m) - \rho^* G(\rho, m)\right) + \delta G(\rho, m)(\rho^* - \tilde{\rho})$$
$$\Rightarrow B\tilde{\rho} = A + \epsilon \left(E(\rho', m') - \rho^* G(\rho', m')\right) + \delta G(\rho, m)(\rho^* - \tilde{\rho})$$
$$\Rightarrow B\tilde{\rho} = A + \epsilon \left(E(\rho', m') - \tilde{\rho} G(\rho', m')\right) + (\delta G(\rho, m) - \epsilon G(\rho', m'))(\rho^* - \tilde{\rho})$$

As $\tilde{\rho} > \rho^*$ and $\delta G(\rho, m) > \epsilon G(\rho', m')$, the final term is negative. We conclude that $\tilde{\rho} < \frac{A + \epsilon E(\rho', m')}{B + \epsilon G(\rho', m')}$. The term on the right is the new posterior for message m; thus, the conditional expected correlation for m has increased.

Therefore, this process effectively splits probability mass from the posterior $\tilde{\rho}$ (generated by m) between ρ^* and a posterior larger than $\tilde{\rho}$. By strict convexity in the neighborhood of ρ^*, this new information structure yields higher profits. Repeating this process until $\int_{M'} \int_{\rho^{**}}^{\bar{\rho}} f(m' \mid \rho') f(\rho') d\rho' dm' = 0$, we conclude that, with probability one, **all correlations $\rho \leq \rho^{**}$ must be pooled to generate the posterior ρ^*.**

Step 5. The convexity of π^{RS} implies that for $\rho > \rho^{**}$, the insurer weakly prefers full revelation of the correlation. This preference is strict wherever the RS solution is interior.

Step 6. The case $\mathbb{E}[\rho] < \rho^*$ is analogous and follows by applying symmetric arguments from Steps 3-5.

□

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.

L. Barseghyan, F. Molinari, T. O'Donoghue, and J. C. Teitelbaum. The nature of risk preferences: Evidence from insurance choices. *American Economic Review*, 103(6):2499–2529, 2013.

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.

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

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

M. K. Brunnermeier, R. Lamba, and C. Segura-Rodriguez. Inverse screening. Princeton University, Cornell University and Banco Central de Costa Rica, 2026a.

M. K. Brunnermeier, R. Lamba, and C. Segura-Rodriguez. *Supplementary Appendix* to inverse selection. Princeton University, Cornell University and Banco Central de Costa Rica, 2026b.

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.

J. Crémer and R. P. McLean. Full extraction of the surplus in Bayesian and dominant strategy auctions. *Econometrica*, 56(6):1247–1257, 1988.

R. Deb, M. Pai, and A.-K. Roesler. How informed do you want your principal to be? University of Toronto and Rice University, 2024.

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, 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.

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.

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

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

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

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 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.

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.

M. H. Riordan and D. E. M. Sappington. Optimal contracts with public ex post information. *Journal of Economic Theory*, 45(1):189–199, 1988.

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.

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