

DISCLOSURE OF CAUSAL VS. CORRELATIONAL INFORMATION IN MARKETS WITH CORRELATION NEGLECT

ANDRIY ZAPECHELNYUK  DIMITRI MIGROW

We study disclosure of consumer information in competitive risk-sharing markets when consumers neglect correlation. Disclosure of a consumer trait mitigates adverse selection, thus improving welfare. It also leads to distorted risk assessment and incorrect participation decisions by consumers due to their lack of understanding of correlation, thus reducing welfare. The net effect on welfare depends on whether the trait's causal or correlational effect on consumer's risk dominates. Stronger causal effect entails more reduction of adverse selection and less distortion of consumers' risk assessment. Thus, our results highlight that causal effects support transparency, while correlational effects may justify privacy protection.

KEYWORDS: information disclosure, adverse selection, correlation neglect, causality, correlation, privacy, insurance market, credit market.

Date: January 26, 2026

Andriy Zapechelnyuk: School of Economics, University of Edinburgh, azapech@gmail.com

Dimitri Migrow: School of Economics, University of Edinburgh, dimitri.migrow@gmail.com

The authors thank Daniel Bird, Andrew Clausen, Inga Deimen, Francesc Dilmé, Federico Echenique, Yingni Guo, Ehud Lehrer, Ludovic Renou, Ran Spiegler, Jakub Steiner, as well as many conference and seminar participants for helpful comments. Author names are in an AEA-certified random order (confirmation code aMD_qj0sTORI).

1. INTRODUCTION

Policies that promote transparency in insurance, credit, and other risk-sharing markets are often motivated by the argument that allowing firms to condition prices on observable consumer characteristics is welfare-improving because it enhances risk segmentation and mitigates adverse selection. However, this argument abstracts from an important asymmetry: while firms and consumers both observe the same characteristics, they may disagree on how these characteristics relate to underlying risk.

A growing empirical literature suggests that firms routinely rely on large datasets and algorithmic methods to estimate how observable traits predict outcomes in a reduced-form sense, while consumers tend to interpret the same traits through causal reasoning based on personal experience and commonly understood causal mechanisms.^{1,2} As a result, consumers may correctly understand how an observable characteristic affects risk through direct causal mechanisms, but fail to account for additional predictive power that arises because the characteristic is correlated with unobserved risk factors that are systematically shared across individuals or persist over time.³ This paper studies the implications of such disagreement for optimal disclosure policy in competitive risk-sharing markets.

A central feature of our analysis is the distinction between the *causal* and the *correlational* components of risk predictiveness. Some observable characteristics affect outcomes primarily through direct causal channels that are widely understood by consumers. Others are predictive mainly because they proxy for unobservable traits or environments—such as income volatility, behavioral stability, or exposure to correlated shocks—that consumers do not observe at the population level. Importantly, the same characteristic may play very

¹For example: (i) in health insurance, smoking causes chronic lung disease; (ii) in life insurance, women live longer than men; (iii) in auto insurance, higher annual mileage is associated with higher risk of accidents. These causal mechanisms are not based on technical expertise or econometric inference, but on a commonsense understanding.

²See Slovic et al. (2016) for an overview of factors determining consumer risk perception. Störmer (2015) uses empirical evidence from auto insurance and life insurance markets to show that consumers understand factors whose risk-determining function is transparent, such as annual mileage in car insurance and smoking in life insurance, but reject person-specific factors, such as homeowner and marital status. Schmeiser et al. (2014) document similar patterns in consumer perceptions of gender as a risk factor across different insurance markets. Kiviat (2019) focuses on the role of credit scores in car insurance pricing and highlights the importance of “...causal understanding and moral categorization for people accepting markets as fair.”

³Some of these risk factors may in principle be observable but are not contractible or legally permissible to use in pricing (e.g., genetic information in health insurance). From the perspective of the pricing problem and the consumer’s model, such factors effectively behave like unobservables.

different roles across markets. For example, gender may be seen as a causal determinant of longevity in annuity markets, but in health or auto insurance it is primarily predictive because it proxies for underlying behavioral, biological, or social factors that are not directly observed or priced.

We formalize this distinction in a competitive market in which a risk-averse consumer faces a lottery whose outcome depends on an observable characteristic and an unobservable shock. Risk-neutral firms compete by offering fixed payments contingent on the observable characteristic. Firms correctly price contracts using the true conditional expectation of the outcome, which incorporates both causal effects and correlations with unobserved residual risk. Consumers, by contrast, understand the direct causal effect of the characteristic but treat the residual risk as independent noise. This difference in interpretation reflects informational constraints rather than irrationality: estimating reduced-form correlations requires access to large datasets, while causal mechanisms can often be inferred from personal experience or established knowledge.

Our focus is on regulation of information disclosure. We say that a characteristic is *disclosed* if firms are allowed to condition prices on it, and *protected* otherwise. When a characteristic is disclosed, firms tailor prices more precisely to consumers' objective risk types. However, if the characteristic is predictive primarily through correlation, consumers may significantly underestimate the total effect of the characteristic on risk. From the consumer's perspective, prices then appear systematically distorted. As a result, consumers may reject contracts that are objectively fair, leading to reduced participation even under perfect competition. Alternatively, when a characteristic is protected, firms must pool consumers across different values of that characteristic. In this case, consumers' beliefs are aligned with pricing. However, pooling induces adverse selection and weakens risk segmentation. The regulator therefore faces a trade-off between reducing adverse selection through disclosure and avoiding participation distortion driven by consumers' lack of understanding correlation.

Our main result shows that disclosure is more likely to improve welfare when an observable characteristic affects outcomes primarily through causal channels and less through correlation with unobserved residual risk. In such cases, consumers' and firms' interpretations of prices are closely aligned, so disclosure alleviates adverse selection without inducing substantial participation distortion. By contrast, when correlational channels dominate, disclosure generates significant disagreement about risk, leading to participation distortion

and, at the same time, limited gains from risk segmentation. In these environments, protecting the characteristic may be welfare-improving despite the presence of adverse selection.

The framework helps rationalize observed heterogeneity in both regulatory outcomes and public acceptance of risk-based pricing. Consumers are more willing to accept price differentiation based on characteristics perceived as causally linked to risk, such as vehicle attributes in auto insurance or health indicators in life insurance, and less willing to accept person-specific characteristics that are viewed as merely correlational, such as homeowner status in both auto and life insurance. Our model provides an explanation for why similar disclosure policies—such as bans on gender-based pricing or community rating requirements—can have sharply different effects across markets.

Related Literature. Our paper contributes to the growing literature on the consequences of behavioral misperceptions for market outcomes and welfare, particularly the effect of correlation neglect by market participants. Correlation neglect has been documented in empirical studies and lab experiments (e.g., [Brunnermeier, 2009](#), [Coval et al., 2009](#), [Hellwig, 2009](#), [Eyster and Weizsacker, 2010](#), [Ortoleva and Snowberg, 2015](#), [Enke and Zimmermann, 2019](#), [Jiao et al., 2020](#)). The theoretical literature studies environments where decision-makers suffer from varying degrees of correlation misperception or misspecification ([DeMarzo et al., 2003](#), [Eyster and Piccione, 2013](#), [Ortoleva and Snowberg, 2015](#), [Levy and Razin, 2015, 2022](#)). [Ellis and Piccione \(2017\)](#) provide an axiomatic foundation for individual choice in an environment where agents fail to account for complexity. Much of this literature focuses on how agents naively combine multiple signals or forecasts (for an overview, see [Levy and Razin 2019](#), Section 3). Our model departs from the literature by applying the concept of correlation neglect to a single-agent inference problem about the data-generating process itself, where the consumer fails to account for the correlation between an observable trait and an unobservable risk factor.

Our work also connects to the literature on economic decision-making under misspecified models, which examines errors in interpreting causal relationships, such as mistakes driven by the reverse-causality error ([Spiegler, 2022](#)) or flawed causal inference from data where decision-makers have a wrong understanding of causality ([Spiegler, 2020, 2025](#)). Another related strand of literature addresses information disclosure in markets ([Bergemann et al., 2015](#), [Azevedo and Gottlieb, 2017](#), [Roesler and Szentes, 2017](#), [Hidir and Velodi, 2021](#), [Garcia and Tsur, 2021](#), [Farinha Luz et al., 2023](#), [Zapechelnyuk and Migrow, 2025](#)). Our paper bridges these strands of literature by embedding a specific form of causal

misperception—correlation neglect in inferring one’s own risk—into a standard competitive market model with disclosure. This allows us to study a new trade-off in disclosure policy that hinges on the causal vs. correlational nature of data, and identify a novel mispricing effect that affects regulatory disclosure decisions in competitive risk-sharing markets.

In a related paper, [Sandroni and Squintani \(2007\)](#) study implications of behavioral biases in competitive insurance. Their regulatory policies and the resulting economic trade-off are distinct from our study, however. Their analysis centers on mandating insurance in the presence of consumer overconfidence. Instead, we analyze effects of transparency (information disclosure) in the presence of consumer correlation neglect. The different policy focus leads us to identify a novel mechanism: the optimal disclosure rule for a consumer characteristic depends on its statistical nature, namely, whether it predicts risk through causal or correlational channels.

2. MODEL

A consumer faces a lottery y determined by real-valued random variables x and ε through a causality equation

$$y = \alpha + \beta x + \varepsilon,$$

where $\alpha \in \mathbb{R}$ and $\beta \geq 0$ are parameters. Variable x is the consumer’s observable characteristic that our analysis focuses on. We will refer to x as the consumer’s *type*. All other observable information is fixed and summarized in the parameter α . Variable ε summarizes unobservable information. W.l.o.g., assume⁴

$$\mathbb{E}[x] = \mathbb{E}[\varepsilon] = 0.$$

The consumer is a risk-averse expected utility maximizer. Her utility from lottery y given type x is $\mathbb{E}[u(y)|x]$, where u is a strictly increasing and strictly concave utility function.

There are several risk-neutral firms that offer nonrandom payments to the consumer in exchange for the lottery y . The firms may or may not observe x (we will analyse and compare both cases). Let I denote what firms observe. The firms compete for the consumer, so the consumer will only trade y for the highest offered payment. Thus, in equilibrium,

⁴This assumption is w.l.o.g., as we can define $y = \hat{\alpha} + \beta \hat{x} + \hat{\varepsilon}$ with $\hat{x} = x - \mathbb{E}[x]$, $\hat{\varepsilon} = \varepsilon - \mathbb{E}[\varepsilon]$, and $\hat{\alpha} = \alpha + \beta \mathbb{E}[x] + \mathbb{E}[\varepsilon]$.

this payment will be equal to the expected value of the lottery given information I :

$$p_I = \mathbb{E}[y|I].$$

The key novelty of our model is that the firms and the consumer have different understanding of the joint distribution of x and ε . The firms believe that x and ε are correlated; specifically, $\mathbb{E}[\varepsilon|x] = \gamma x$, where $\gamma \geq 0$ is a parameter. Thus, the random payoff y is affected by x directly (with weight β) and indirectly through correlation with ε (with weight γ). Unlike the firms, the consumer suffers from correlation neglect and thinks that x and ε are independent. For tractability, assume that from the firms' perspective ε can be written as the sum of its expected value γx and independent noise z with zero mean:

$$\varepsilon = \gamma x + z, \text{ with } \mathbb{E}[z] = 0.$$

From the consumer's perspective, ε is pure noise, $\varepsilon = z$. Substituting ε with the respective interpretations of the firms and the consumer, we obtain:

$$\text{Firms' model: } y = \alpha + (\beta + \gamma)x + z.$$

$$\text{Consumer's model: } y = \alpha + \beta x + z.$$

Assume that x has convex support $X \subseteq \mathbb{R}$ and density $f(x)$, and z has convex support $Z \subseteq \mathbb{R}$ and density $g(z)$. In what follows, for any given function $\phi(x, z)$, we use the notation

$$\mathbb{E}[\phi(x, z)|x] = \int_Z \phi(x, z)g(z)dz \text{ and } \mathbb{E}[\phi(x, z)] = \int_{X \times Z} \phi(x, z)g(z)f(x)dzdx.$$

We evaluate the welfare from ex-ante perspective (before x is realized) under the true (firms') model. We interpret the consumer as being mistaken by disregarding the existing correlation between x and ε . Since the market is competitive and the firms make no profit, the welfare is evaluated as the consumer surplus.

3. ANALYSIS

In this section, we first analyze two cases: x is disclosed (the firms can condition their payments on x) and x is protected (the firms must offer the same payment for all x), where

we use subscripts D and P to refer to disclosure and protection, respectively. We then compare these cases and provide a comparative statics.

For the purpose of comparative statics, we consider a different parametrization. Assume that $\beta + \gamma > 0$, and let

$$c = \beta + \gamma \quad \text{and} \quad \lambda = \frac{\beta}{\beta + \gamma}.$$

Parameter c reflects the firms' coefficient attached to variable x . In practice, this is the coefficient that the firms estimate from data. Parameter λ captures the degree of *causality* of x . The higher λ , the more x directly causes y and the less it affects y through correlation with ε . So, x is a pure causality variable when $\lambda = 1$ and it is a pure correlation variable when $\lambda = 0$. Solving for β and γ as functions of c and λ , we obtain:

$$\text{Firms' model: } y = \alpha + cx + z,$$

$$\text{Consumer's model: } y = \alpha + \lambda cx + z.$$

For the purpose of comparative statics, we will vary λ in $[0, 1]$ while keeping c constant. Thus, we maintain the true informational value of variable x for understanding the risks from the firms' perspective and only vary the degree of causality of x .

3.1. Variable x is disclosed

Suppose that x is disclosed, so $I = x$. As the firms can condition their payments on x ,

$$p_x = \mathbb{E}[y|x] = \mathbb{E}[\alpha + cx + z|x] = \alpha + cx.$$

Given type x , the consumer obtains the utility $u(p_x) = u(\alpha + cx)$ by accepting the payment and the expected utility $\mathbb{E}[u(\alpha + \lambda cx + z)|x]$ by rejecting it. Let $X_D^*(\lambda)$ be the subset of values of x such that the consumer accepts payment $p_x = \alpha + cx$:

$$X_D^*(\lambda) = \{x \in X : u(\alpha + cx) \geq \mathbb{E}[u(\alpha + \lambda cx + z)|x]\}. \quad (1)$$

Recall that the welfare is the consumer surplus calculated according to the true (firms') model. Here it is given by

$$W_D(\lambda) = \int_{x \in X_D^*(\lambda)} u(\alpha + cx)f(x)dx + \int_{x \notin X_D^*(\lambda)} \mathbb{E}[u(\alpha + cx + z)|x]f(x)dx. \quad (2)$$

The next proposition shows that, when x is disclosed, as λ increases (so x becomes more causal), more types of consumers accept the payment and the welfare strictly increases. Moreover, when λ is above some threshold λ_1 , all types of consumers accept the payment and the market is efficient.

PROPOSITION 1: *There exists $\lambda_1 \in [0, 1]$ such that*

- (i) $X_D^*(\lambda) = X$ and $W_D(\lambda) = \mathbb{E}[u(\alpha + cx)]$ for all $\lambda \geq \lambda_1$, and
- (ii) $X_D^*(\lambda)$ and $W_D(\lambda)$ are strictly increasing⁵ in λ for all $\lambda < \lambda_1$.

The proof of this and other propositions are deferred to the Appendix.

3.2. Variable x is protected

Suppose that x is protected, so $I = \emptyset$. As the firms are not allowed to condition their payments on x , they must give the same payment p to all types x . Denote

$$\underline{x} = \inf X \text{ and } \bar{x} = \sup X.$$

Note that \underline{x} and \bar{x} may be infinite if X is unbounded.

Given a payment p and a type x , the consumer gets the utility $u(p)$ by accepting the payment and the expected utility $\mathbb{E}[u(\alpha + \lambda cx + z)|x]$ by rejecting it. When $\lambda > 0$, the expected utility is strictly increasing in x . Thus, there exists a participation threshold x^* such that consumers with types $x \leq x^*$ accept p and consumers with types $x > x^*$ reject it. In particular, $x^* = \bar{x}$ if all types accept the payment and $x^* = \underline{x}$ if all reject the payment. When $\lambda = 0$, the expected utility $\mathbb{E}[u(\alpha + z)|x]$ is independent of x , and thus either all types accept the payment ($x^* = \bar{x}$) or all reject the payment ($x^* = \underline{x}$).

By zero profit, in equilibrium, the payment must satisfy

$$p = \mathbb{E}[\alpha + cx + z|x \leq x^*] = \alpha + c\mathbb{E}[x|x \leq x^*].$$

⁵Here and below, set $X_D^*(\lambda)$ is increasing in λ if, for all $\lambda' < \lambda''$, $X_D^*(\lambda') \subseteq X_D^*(\lambda'')$; moreover, it is strictly increasing if the set $X_D^*(\lambda'') \setminus X_D^*(\lambda')$ has nonempty interior.

Denote by $\Delta(x^*, \lambda)$ the difference of the utility of consumer with type x^* from accepting and rejecting the payment $p = \alpha + c\mathbb{E}[x|x \leq x^*]$:

$$\Delta(x^*, \lambda) = u(\alpha + c\mathbb{E}[x|x \leq x^*]) - \mathbb{E}[u(\alpha + \lambda cx^* + z)|x^*]. \quad (3)$$

Due to free entry, the equilibrium participation threshold is the highest x^* such that consumers with types $x \leq x^*$ accept $p = \alpha + c\mathbb{E}[x|x \leq x^*]$, and there does not exist a higher payment with the same property:

$$x_P^*(\lambda) = \max \left\{ x^* \in X : \Delta(x^*, \lambda) \geq 0 \right\}, \quad (4)$$

with the convention that $x_P^*(\lambda) = \underline{x}$ if $\Delta(x^*, \lambda) < 0$ for all $x^* \in X$, and $x_P^*(\lambda) = \bar{x}$ if $\Delta(x^*, \lambda) \geq 0$ for all $x^* \in X$.

Note that (4) is conceptually identical to the standard adverse selection problem in a competitive market setting, with the only difference in a distortion of the consumer's participation decision when $\lambda < 1$. To obtain a clean comparative statics with respect to λ , some of our results will require the following assumption:

$$x_P^*(\lambda) \text{ is continuous in } \lambda. \quad (\text{A}_1)$$

Loosely speaking, this assumption holds if the primitives of the problem are “regular” enough to prevent discontinuous jumps of the equilibrium in response to small changes of λ . We further discuss this assumption in Section 4.

Recall that the welfare is the consumer surplus calculated according to the true (firms') model. Here it is given by

$$\begin{aligned} W_P(\lambda) &= u(\alpha + c\mathbb{E}[x|x \leq x_P^*(\lambda)]) \int_{x \leq x_P^*(\lambda)} f(x) dx \\ &\quad + \int_{x > x_P^*(\lambda)} \mathbb{E}[u(\alpha + cx + z)|x] f(x) dx. \end{aligned} \quad (5)$$

The next proposition shows that, when x is protected, as λ increases (so x becomes more causal), fewer types of consumers accept the payment and the welfare weakly decreases. Moreover, when λ is below some threshold λ_0 , all types of consumers accept the payment and the market is efficient.

PROPOSITION 2: *There exists $\lambda_0 \in [0, 1]$ such that*

(i) $x_P^(\lambda) = \bar{x}$ and $W_P(\lambda) = u(\alpha)$ for all $\lambda \leq \lambda_0$, and*

(ii) if Assumption (A₁) holds, then $x_P^(\lambda)$ and $W_P(\lambda)$ are decreasing in λ .*

3.3. Comparative Statics

We now evaluate the difference in welfare between disclosed and protected x , and analyze how this difference changes as λ goes up, that is, as variable x becomes more causal.

Let us make two observations. First, consider the case where x is protected, and suppose that the consumer is forced to accept the contract (mandatory participation). Then, adverse selection is absent and the equilibrium payment is $p^* = \mathbb{E}[\alpha + cx + z] = \alpha$. Since there is mandatory participation, the welfare is $u(\mathbb{E}[\alpha + cx + z]) = u(\alpha)$.

Second, consider the case where x is disclosed and the consumer has no correlation neglect, so $\lambda = 1$. Then, for each type x , the equilibrium payment is $p_x = \mathbb{E}[\alpha + cx + z|x] = \alpha + cx$, and the consumer prefers to accept this payment. Thus, the welfare is $\mathbb{E}[u(\alpha + cx)]$.

Given the above observations, write

$$W_D(\lambda) - W_P(\lambda) = \left(W_D(\lambda) - \mathbb{E}[u(\alpha + cx)] \right) + \left(\mathbb{E}[u(\alpha + cx)] - u(\alpha) \right) + \left(u(\alpha) - W_P(\lambda) \right).$$

The part $\left(W_D(\lambda) - \mathbb{E}[u(\alpha + cx)] \right)$ is the *correlation neglect effect* as it compares the surplus when x is disclosed with and without correlation neglect of the consumer. This effect is negative, and it vanishes when $\lambda = 1$.

The part $\left(\mathbb{E}[u(\alpha + cx)] - u(\alpha) \right)$ is known in the literature as the *price uncertainty effect*, as it compares the surplus with and without disclosure of x under full participation. This effect is always negative.⁶

The part $\left(u(\alpha) - W_P(\lambda) \right)$ is the *adverse selection effect*, as it compares the surplus under nondisclosure of x with and without mandatory participation. This effect is positive, and it vanishes when $\lambda = 0$.

As λ goes up, so variable x affects y more through the causality channel and less through the correlation channel, the correlation neglect effect diminishes by Proposition 1, the adverse selection effect strengthens by Proposition 2 with Assumption (A₁), and the price

⁶The literature on price uncertainty goes back to [Hirshleifer \(1971\)](#); for recent papers, see [Farinha Luz et al. \(2023\)](#) and [Veiga \(2024\)](#).

uncertainty effect is not affected by λ . It follows that $W_D(\lambda) - W_P(\lambda)$ is increasing in λ , so the more causal x is, the more reason to disclose it.

When Assumption (A₁) does not hold, we cannot establish a monotone comparative statics of the adverse selection effect in λ . Yet, we can compare the extreme cases.

Suppose that λ is small, namely, $\lambda \leq \lambda_0$, where λ_0 is the constant in Proposition 2. In words, variable x affects y predominantly through the correlation channel, which is neglected by the consumers. Then, by Proposition 2, the adverse selection effect is absent, whereas, by Proposition 1, the correlation neglect is maximal. It is then unambiguous that $W_D(\lambda) < W_P(\lambda)$, so it is optimal to protect x .

Now, suppose that λ is large, namely, $\lambda \geq \lambda_1$, where λ_1 is the constant in Proposition 1. In words, variable x affects y predominantly through the causality channel. In this case, the consumers' and the firms' evaluation of y are closely aligned, and, by Proposition 1, the correlation neglect is absent. This becomes the standard adverse selection story without any novel elements. Whether it is better to disclose or to protect x depends on whether adverse selection dominates the price uncertainty effect, which is a well known tradeoff in the literature.

4. DISCUSSION

Competition. We assume that the market is competitive. If we relax this assumption, then a new reason to protect x appears: the protection of the consumer's information rent. Nevertheless, the tradeoff exposed in our analysis that emerges due to the consumer's correlation neglect remains valid and relevant in a setting of oligopolistic or monopolistic markets.

Parameter λ . We assume that the parameter λ that captures the degree of causality of x is in $[0, 1]$. However, the model remains valid if λ is allowed to be in \mathbb{R} . This means that coefficients β and γ may be positive or negative, provided $\beta + \gamma \neq 0$.

W.l.o.g. assume that $c = \beta + \gamma > 0$ (otherwise, change the sign of x). When $\lambda < 0$, the consumer believes that x affects y negatively, but the correlation of x and ε is so strong that it flips the sign of the relationship for the firms. Qualitatively, this case is very similar to the case of $\lambda = 0$. There is no adverse selection, and thus, disclosure of x brings no benefit. So, welfare is maximized by protecting x .

When $\lambda > 1$, the consumer believes that x affects y more than it actually does. The correlation neglect effect is absent if λ is close to 1, but it emerges and becomes stronger as λ increases. Under Assumption (A₁), the adverse selection also becomes stronger as λ

increases. The net effect—whether disclosure of x is desirable—is ambiguous for a range of values of λ . But when λ is large enough, the correlation neglect effect eventually dominates, and welfare is maximized by protecting x .

Assumption (A₁). It is difficult to provide an economic interpretation to Assumption (A₁), as it is a technical condition on an endogenous object, $x_P^*(\lambda)$. In this section, we provide a different sufficient condition for part (ii) of Proposition 2 and discuss when this condition is satisfied.

PROPOSITION 2': *If*

$$u(\alpha + c\mathbb{E}[x|x \leq 0]) \geq \mathbb{E}[u(\alpha + z)|x = 0], \quad (\text{A}_2)$$

then $x_P^(\lambda)$ and $W_P(\lambda)$ are decreasing in λ .*

To understand assumption (A₂), consider the traditional adverse selection setting, which, in our model, corresponds to $\lambda = 1$ and x protected. Then, (A₂) means that the firms do not make losses by offering payment $p = \mathbb{E}[\alpha + cx + z|x \leq 0] = \alpha + c\mathbb{E}[x|x \leq 0]$. Indeed, if p is offered, (A₂) states that the consumer with type $x = 0$ weakly prefers to accept the payment p , and thus, all consumers with types below 0 also prefer to accept p . The immediate implication is that the equilibrium threshold type, $x_P^*(\lambda)$ with $\lambda = 1$, is weakly above the mean type, $x = 0$. This can be interpreted as the condition that the adverse selection is not too severe: at least the types below the mean type participate in equilibrium.

We now discuss in more detail how the model primitives (namely, utility u and densities f and g) affect whether assumption (A₂) holds or not. Let

$$\begin{aligned} \Phi(u, f, g) &= \alpha + c\mathbb{E}[x|x \leq 0] - u^{-1}(\mathbb{E}[u(\alpha + z)]) \\ &= \alpha + \frac{c}{F(0)} \int_{x \leq 0} xf(x)dx - u^{-1}\left(\int_Z u(\alpha + z)g(z)dz\right), \end{aligned} \quad (6)$$

where F denotes the CDF of f . Note that (A₂) holds if and only if $\Phi(u, f, g) \geq 0$.

For two strictly increasing and strictly concave functions u_1 and u_2 , say that u_1 is *more concave than* u_2 if $u_1(u_2^{-1})$ is concave.⁷

PROPOSITION 3:

- (i) If u_1 is more concave than u_2 , then $\Phi(u_1, f, g) \geq \Phi(u_2, f, g)$.
- (ii) If g_1 is a mean-preserving spread of g_2 , then $\Phi(u, f, g_1) \geq \Phi(u, f, g_2)$.
- (iii) If f_1 is a mean-preserving spread of f_2 and $F_1(0) \leq F_2(0)$, then $\Phi(u, f_1, g) \leq \Phi(u, f_2, g)$.

The implication of Proposition 3 is that primitives (u, f, g) do not satisfy (A₂) if either the utility u is not concave enough, or the distribution of z is not spread out enough, or, loosely speaking, the distribution of x is too spread out to the left of its mean. For example, if u belongs to the constant relative risk aversion (CARA) family, $u(x) = -e^{-kx}$ with risk aversion parameter $k > 0$, and x and z are normally distributed with zero mean and standard deviations σ_x and σ_z , respectively, then (A₂) simplifies to

$$\frac{k\sigma_z^2}{\sigma_x} \geq \sqrt{\frac{2}{\pi}}.$$

In words, (A₂) holds if the consumer is sufficiently risk averse (k is high enough), the noise variable z has a high enough variance, and the observable variable x has a low enough variance. The above inequality also shows the relative strength of different effects, for example, the variance of z is more impactful than the variance of x .

⁷Note that when u_1 and u_2 are strictly increasing and twice differentiable, $u_1(u_2^{-1})$ is concave if and only if $u'_1(x)/u'_2(x)$ is decreasing, or equivalently, the Arrow-Pratt measure of risk aversion is uniformly higher for u_1 than for u_2 , that is, $-u''_1(x)/u'_1(x) \geq -u''_2(x)/u'_2(x)$ for all x .

Proof of Proposition 1. Let

$$\Lambda_1 = \left\{ \lambda \in [0, 1] : u(\alpha + cx) \geq \mathbb{E}[u(\alpha + \lambda cx + z)|x] \text{ for all } x \in X \right\}.$$

Let λ_1 be the smallest element of Λ_1 . Note that λ_1 is well defined, since Λ_1 is closed (by continuity of u), and $1 \in \Lambda_1$. Indeed, if $\lambda = 1$, then we have $u(\alpha + cx) \geq \mathbb{E}[u(\alpha + cx + z)]$ by Jensen's inequality and $\mathbb{E}[z] = 0$.

Part (i). Consider $\lambda \geq \lambda_1$. By (1), (2), and definition of λ_1 , we have $X_D^*(\lambda) = X$ and $W_D(\lambda) = \mathbb{E}[u(\alpha + cx)]$ for all $\lambda \geq \lambda_1$.

Part (ii). First, for each $x \geq 0$ we have

$$u(\alpha + cx) - \mathbb{E}[u(\alpha + \lambda cx + z)|x] \geq u(\alpha + cx) - u(\alpha + \lambda cx) \geq 0,$$

where the first inequality is by Jensen's inequality and $\mathbb{E}[z] = 0$, and the second inequality is because u is increasing, $c > 0$, and $x \geq 0$. Thus, by (1), we have:

$$\text{If } x \geq 0, \text{ then } x \in X_D^*(\lambda) \text{ for all } \lambda. \quad (7)$$

Next, because u is strictly increasing and $c > 0$, we have

$$\text{If } x < 0, \text{ then } \mathbb{E}[u(\alpha + \lambda cx + z)|x] \text{ is strictly decreasing in } \lambda. \quad (8)$$

Next, consider $\lambda', \lambda'' \in [0, 1]$ such that $\lambda' < \lambda'' \leq \lambda_1$. By (1) and the continuity of u , the set $X_D^*(\lambda')$ is closed. By $\lambda' < \lambda_1$ and definition of λ_1 , the set $X \setminus X_D^*(\lambda')$ is nonempty. Hence, $X \setminus X_D^*(\lambda')$ has nonempty interior. It follows from (1), (7), and (8) that $X_D^*(\lambda') \subsetneq X_D^*(\lambda'')$, and moreover, $X_D^*(\lambda'') \setminus X_D^*(\lambda')$ has nonempty interior. Finally, we have

$$W_D(\lambda'') - W_D(\lambda') = \int_{x \in X_D^*(\lambda'') \setminus X_D^*(\lambda')} (u(\alpha + cx) - \mathbb{E}[u(\alpha + cx + z)|x]) f(x) dx > 0,$$

where the equality is by (2) and $X_D^*(\lambda') \subsetneq X_D^*(\lambda'')$, and the inequality is because u is strictly concave, $\mathbb{E}[z] = 0$, and $X_D^*(\lambda'') \setminus X_D^*(\lambda')$ has a positive measure under f . *Q.E.D.*

Auxiliary Lemma. The next lemma will be used in the proof of Propositions 2 and 2'.

LEMMA 1: Suppose that $x_P^*(\lambda) \geq 0$ for all $\lambda \in [0, 1]$. Then $x_P^*(\lambda)$ and $W_P(\lambda)$ are decreasing in λ .

PROOF: First, we prove that $x_P^*(\lambda)$ is decreasing in λ . Consider $\lambda', \lambda'' \in [0, 1]$ such that $\lambda'' > \lambda' \geq \lambda_0$. We have

$$\begin{aligned} \Delta(x^*, \lambda'') &= u(\alpha + c\mathbb{E}[x|x \leq x^*]) - \mathbb{E}[u(\alpha + \lambda''cx^* + z)|x^*]] \\ &< u(\alpha + c\mathbb{E}[x|x \leq x^*]) - \mathbb{E}[u(\alpha + \lambda'cx^* + z)|x^*]] \\ &= \Delta(x^*, \lambda') < 0, \quad \text{for all } x^* > x_P^*(\lambda'). \end{aligned} \quad (9)$$

Here, both equalities are by (3). The first inequality is because $\lambda'' > \lambda'$, $c > 0$, $x^* > 0$ (since $x^* > x_P^*(\lambda')$ and, by the Lemma assumption, $x_P^*(\lambda') \geq 0$), and u is strictly increasing. The second inequality is because $x_P^*(\lambda')$ satisfies (4) and $x^* > x_P^*(\lambda')$. It follows from (4) and (9) that $x_P^*(\lambda'') \leq x_P^*(\lambda')$.

Next, we prove that $W_P(\lambda)$ is decreasing in λ . Let

$$W(x^*) = u(\alpha + c\mathbb{E}[x|x \leq x^*]) \int_{x \leq x^*} f(x)dx + \int_{x > x^*} \mathbb{E}[u(\alpha + cx + z)|x]f(x)dx. \quad (10)$$

For all $x^* \geq 0$, we have

$$\begin{aligned} \frac{d}{dx^*} W(x^*) &= (\mathbb{E}[u(\alpha + cx^* + z)|x^*] - u(\alpha + c\mathbb{E}[x|x \leq x^*])) f(x^*) \\ &\quad + cu'(\alpha + c\mathbb{E}[x|x \leq x^*]) \frac{d\mathbb{E}[x|x \leq x^*]}{dx^*} \int_{x \leq x^*} f(x)dx \\ &\geq (\mathbb{E}[u(\alpha + \lambda cx^* + z)|x^*] - u(\alpha + c\mathbb{E}[x|x \leq x^*])) f(x^*) \\ &= -\Delta(x^*, \lambda) f(x^*), \end{aligned} \quad (11)$$

Here, the first equality is by differentiation of (10), the inequality is because u is increasing, $c > 0$, $x^* \geq 0$, $\lambda \leq 1$, and $d\mathbb{E}[x|x \leq x^*]/dx^* \geq 0$, and the last equality is by (3). Thus, for almost all $\lambda > \lambda_0$ we obtain⁸

⁸As $x_P^*(\lambda)$ is monotone, it is almost everywhere differentiable (Lebesgue's Theorem).

$$\frac{dW_P(\lambda)}{d\lambda} = \left. \frac{dW(x^*)}{dx^*} \right|_{x^*=x_P^*(\lambda)} \times \frac{dx_P^*(\lambda)}{d\lambda} \leq -\Delta(x_P^*(\lambda), \lambda) f(x_P^*(\lambda)) \frac{dx_P^*(\lambda)}{d\lambda} = 0.$$

Here, the first equality is by (5) and (10), the inequality is by (11) and because $x_P^*(\lambda)$ is decreasing in λ (as proved above), and the last equality is because $x_P^*(\lambda)$ satisfies (4) and $\lambda > \lambda_0$, so $\Delta(x_P^*(\lambda), \lambda) = 0$. *Q.E.D.*

Proof of Proposition 2. Let

$$\Lambda_0 = \left\{ \lambda \in [0, 1] : u(\alpha) \geq \mathbb{E}[u(\alpha + \lambda cx + z)|x] \text{ for all } x \in X \right\}.$$

Let λ_0 be the greatest element of Λ_0 . Note that λ_0 is well defined, since Λ_0 is closed (by continuity of u), and $0 \in \Lambda_0$. Indeed, if $\lambda = 0$, then we have $u(\alpha) \geq \mathbb{E}[u(\alpha + z)]$ by Jensen's inequality and $\mathbb{E}[z] = 0$.

Part (i). Consider $\lambda \leq \lambda_0$. By (4) and definition of λ_0 , we have $x_P^*(\lambda) = \bar{x}$, and thus, $\mathbb{E}[x|x \leq x_P^*(\lambda)] = \mathbb{E}[x] = 0$. Therefore, by (5), we have $W_P(\lambda) = u(\alpha)$ for all $\lambda \leq \lambda_0$.

Part (ii). It suffices to show that if Assumption (A₁) holds, then $x_P^*(\lambda) \geq 0$ for all $\lambda \in [0, 1]$. Then, part (ii) of Proposition 2 is immediate by Lemma 1.

Suppose that (A₁) holds. By part (i) of Proposition 2, when $\lambda = 0$, we have $x_P^*(0) = \bar{x} > 0$. Suppose by contradiction that $x_P^*(\lambda') < 0$ for some $\lambda' > 0$. Then, by (A₁) and the Intermediate Value Theorem, there exists $\lambda'' \in (0, \lambda']$ such that $x_P^*(\lambda'') = 0$. Also, since $x_P^*(\lambda')$ and $x_P^*(\lambda'')$ satisfy (4) under λ' and λ'' , respectively, we must have $\Delta(0, \lambda') < 0 = \Delta(0, \lambda'')$. However, by (3), the expression

$$\Delta(0, \lambda) = u(\alpha + c\mathbb{E}[x|x \leq x^*]) - \mathbb{E}[u(\alpha + z)|x^*]$$

is independent of λ , which contradicts $\Delta(0, \lambda') < 0 = \Delta(0, \lambda'')$. Thus, we conclude that $x_P^*(\lambda) \geq 0$ for all $\lambda \in [0, 1]$. *Q.E.D.*

Proof of Proposition 2'. By (3) and Assumption (A₂), we have $\Delta(0, \lambda) \geq 0$ for all $\lambda \in [0, 1]$. It follows from (4) that $x_P^*(\lambda) \geq 0$ for all $\lambda \in [0, 1]$. Then, by Lemma 1, we obtain that $x_P^*(\lambda)$ and $W_P(\lambda)$ are decreasing in λ . *Q.E.D.*

Proof of Proposition 3. *Part (i).* Since $u_1(u_2^{-1})$ is concave, by Jensen's inequality, we have

$$u_1(u_2^{-1}(\mathbb{E}[u_2(\alpha + z)])) \geq \mathbb{E}[u_1(u_2^{-1}(u_2(\alpha + z)))] = \mathbb{E}[u_1(\alpha + z)].$$

Applying u_1^{-1} to both sides of the above inequality, we obtain

$$u_2^{-1}(\mathbb{E}[u_2(\alpha + z)]) \geq u_1^{-1}(\mathbb{E}[u_1(\alpha + z)]),$$

which, by (6), is equivalent to $\Phi(u_1, f, g) \geq \Phi(u_2, f, g)$.

Part (ii). Since u is a concave function, if g_1 is a mean-preserving spread of g_2 , then

$$\int_Z u(\alpha + z)g_1(z)dz \leq \int_Z u(\alpha + z)g_2(z)dz,$$

which, by (6) and that the fact that u^{-1} is increasing, is equivalent to $\Phi(u, f, g_1) \geq \Phi(u, f, g_2)$.

Part (iii). Since $\min\{x, 0\}$ is a concave function, if f_1 is a mean-preserving spread of f_2 , then

$$\int_{x \leq 0} x f_1(x)dx = \int_{x \in X} \min\{x, 0\} f_1(x)dx \leq \int_{x \in X} \min\{x, 0\} f_2(x)dx = \int_{x \leq 0} x f_2(x)dx.$$

Note that $F_i(0) > 0$ for each $i = 1, 2$, since F_i is a CDF (and, thus, right-continuous) and $\mathbb{E}[x]$ under F_i is equal to zero. Thus, since $F_2(0) \geq F_1(0) > 0$, we obtain

$$\frac{1}{F_1(0)} \left(- \int_{x \leq 0} x f_1(x)dx \right) \geq \frac{1}{F_2(0)} \left(- \int_{x \leq 0} x f_2(x)dx \right),$$

which, by (6), is equivalent to $\Phi(u, f_1, g) \leq \Phi(u, f_2, g)$.

Q.E.D.

REFERENCES

- AZEVEDO, EDUARDO M AND DANIEL GOTTLIEB (2017): “Perfect competition in markets with adverse selection,” *Econometrica*, 85 (1), 67–105. [4]
- BERGEMANN, DIRK, BENJAMIN BROOKS, AND STEPHEN MORRIS (2015): “The limits of price discrimination,” *American Economic Review*, 105 (3), 921–957. [4]
- BRUNNERMEIER, MARKUS K (2009): “Deciphering the liquidity and credit crunch 2007–2008,” *Journal of Economic perspectives*, 23, 77–100. [4]
- COVAL, JOSHUA, JAKUB JUREK, AND ERIK STAFFORD (2009): “The economics of structured finance,” *Journal of Economic Perspectives*, 23, 3–25. [4]
- DEMARZO, PETER M, DIMITRI VAYANOS, AND JEFFREY ZWIEBEL (2003): “Persuasion bias, social influence, and unidimensional opinions,” *The Quarterly journal of economics*, 118 (3), 909–968. [4]
- ELLIS, ANDREW AND MICHELE PICCIONE (2017): “Correlation misperception in choice,” *American Economic Review*, 107 (4), 1264–1292. [4]
- ENKE, BENJAMIN AND FLORIAN ZIMMERMANN (2019): “Correlation neglect in belief formation,” *The Review of Economic Studies*, 86 (1), 313–332. [4]
- EYSTER, ERIK AND MICHELE PICCIONE (2013): “An approach to asset pricing under incomplete and diverse perceptions,” *Econometrica*, 81, 1483–1506. [4]
- EYSTER, ERIK AND GEORG WEIZSACKER (2010): “Correlation neglect in financial decision-making,” *DIW Berlin Discussion Paper*. [4]
- FARINHA LUZ, VITOR, PIERO GOTTARDI, AND HUMBERTO MOREIRA (2023): “Risk Classification in Insurance Markets with Risk and Preference Heterogeneity,” *Review of Economic Studies*, 90, 3022–3082. [4, 10]
- GARCIA, DANIEL AND MATAN TSUR (2021): “Information design in competitive insurance markets,” *Journal of Economic Theory*, 191, 105160. [4]
- HELLWIG, MARTIN F (2009): “Systemic risk in the financial sector: An analysis of the subprime-mortgage financial crisis,” *De Economist*, 157, 129–207. [4]
- HIDIR, SINEM AND NIKHIL VELLODI (2021): “Privacy, personalization, and price discrimination,” *Journal of the European Economic Association*, 19 (2), 1342–1363. [4]
- HIRSHLEIFER, JACK (1971): “The private and social value of information and the reward to inventive activity,” *American Economic Review*, 61 (4), 561–574. [10]
- JIAO, PEIRAN, ANDRE VEIGA, AND ANSGAR WALTHER (2020): “Social media, news media and the stock market,” *Journal of Economic Behavior & Organization*, 176, 63–90. [4]
- KIVIAT, BARBARA (2019): “The moral limits of predictive practices: The case of credit-based insurance scores,” *American sociological review*, 84, 1134–1158. [2]
- LEVY, GILAT AND RONNY RAZIN (2015): “Correlation neglect, voting behavior, and information aggregation,” *American Economic Review*, 105, 1634–1645. [4]
- (2019): “Echo chambers and their effects on economic and political outcomes,” *Annual Review of Economics*, 11 (1), 303–328. [4]
- (2022): “Combining forecasts in the presence of ambiguity over correlation structures,” *Journal of Economic Theory*, 199, 105075. [4]

- ORTOLEVA, PIETRO AND ERIK SNOWBERG (2015): “Overconfidence in political behavior,” *American Economic Review*, 105 (2), 504–535. [4]
- ROESLER, ANNE-KATRIN AND BALÁZS SZENTES (2017): “Buyer-optimal learning and monopoly pricing,” *American Economic Review*, 107 (7), 2072–2080. [4]
- SANDRONI, ALVARO AND FRANCESCO SQUINTANI (2007): “Overconfidence, Insurance and Paternalism,” *The American Economic Review*, 97 (5), 1994–2004. [5]
- SCHMEISER, HATO, TINA STÖRMER, AND JOËL WAGNER (2014): “Unisex insurance pricing: Consumers’ perception and market implications,” *The Geneva Papers on Risk and Insurance-Issues and Practice*, 39 (2), 322–350. [2]
- SLOVIC, PAUL, BARUCH FISCHHOFF, AND SARAH LICHTENSTEIN (2016): “Facts and fears: Understanding perceived risk,” in *The Perception of Risk*, Routledge, 137–153. [2]
- SPIEGLER, RAN (2020): “Behavioral implications of causal misperceptions,” *Annual Review of Economics*, 12 (1), 81–106. [4]
- (2022): “On the behavioral consequences of reverse causality,” *European Economic Review*, 149, 104258. [4]
- (2025): “Behavioral causal inference,” *Review of Economic Studies*, forthcoming. [4]
- STÖRMER, TINA (2015): “Optimizing insurance pricing by incorporating consumers’ perceptions of risk classification,” *Zeitschrift für die gesamte Versicherungswissenschaft*, 104 (1), 11–37. [2]
- VEIGA, ANDRE (2024): “Price Discrimination in Selection Markets,” *Review of Economics and Statistics*, forthcoming. [10]
- ZAPECHELNYUK, ANDRIY AND DIMITRI MIGROW (2025): “Disclosure in insurance markets with limited screening,” *Working paper*. [4]