Adverse selection, efficiency and the structure of information

Heski Bar-Isaac¹ · Ian Jewitt² · Clare Leaver³

Received: 29 August 2019 / Accepted: 17 July 2020 / Published online: 18 August 2020
© The Author(s) 2020

Abstract
This paper explores how the structure of asymmetric information impacts on economic outcomes in Akerlof’s (Q J Econ 84(3):488–500, 1970) Lemons model applied to the labour market and extended to admit a matching component between worker and firm. We characterize the nature of equilibrium and define measures of adverse selection and efficiency. We then characterize the joint distribution of outcomes—adverse selection, probability of trade, efficiency, profits, and wage—for the class of Gaussian basic games and information, and perform comparative statics with respect to a parsimonious parameterization of the information structure. We use this framework to revisit the classic issue, first addressed by Roy (Oxford Econ Pap 3(2):135-146, 1951), of selection into different sectors. We identify conditions under which an effect reversal—adverse selection at any realisation of public information but, overall, positive selection into the outside sector—can and cannot arise, and note the implications for empirical work. We also explore the divisions of expected total surplus between worker and firm that can be achieved as information varies. We show that, if the distribution of worker types is non-singular, any point in the set of possible surplus divisions can be achieved as a limit of a PBE for some information structure with asymmetric information. Finally, re-interpreting the model in an insurance context, where the matching component becomes consumer risk aversion, we use our framework to highlight sources of advantageous selection.

We are grateful to very many people for helpful comments on earlier versions of this paper, in particular Ignacio Esponda, Paul Klemperer, Jonathan Levin, Alessandro Lizzeri, Margaret Meyer and Dan Quigley. We also thank participants at numerous conferences and seminars. Bar-Isaac thanks SSHRC (435-2014-0004) for financial support. Jewitt and Leaver are grateful for the hospitality of the Toulouse School of Economics, 2018–2019.

Ian Jewitt
ian.jewitt@nuffield.ox.ac.uk

¹ University of Toronto, CEPR and CRESSE, Toronto, Canada
² Nuffield College, University of Oxford and CEPR, Oxford, England
³ Blavatnik School of Government, University of Oxford and CEPR, Oxford, England
Keywords  Information design · Adverse selection · Asymmetric information

JEL Classification  D82 · J30

1 Introduction

We aim to characterize how the structure of asymmetric information impacts economic outcomes in a natural generalization of Akerlof’s (1970) Lemons model.

We couch our discussion in terms of an employer learning model of the labour market, giving a brief discussion of insurance. This model generalizes Akerlof (1970) in two main respects. First, the scalar type assumption is relaxed by introducing a match component between the good to be traded (a worker) and the seller (the current employer, or inside firm). The worker’s productivity in any employment equals some underlying general productivity plus a random match value. Second, we admit general information structures consisting of both public and private information. The inside firm is assumed to privately observe a multidimensional statistic \( Q \) that carries information about the worker’s general productivity and about how well she is matched in her current employment. Outside firms also observe information about the worker; public information that we denote by the random vector \( T \). Asymmetric information means that given \( Q, T \) is assumed to be conditionally independent of productivity and match value. A description of the model is provided in Sect. 2.

We begin our analysis in Sect. 3 by considering the wage determination process. Our model of competition for the worker follows the classic paper by Greenwald (1986). The inside firm faces Bertrand competition from a set of identical outside firms who make wage offers that, if successful, result in zero expected profit. Our first result establishes conditions (on the information structure) under which perfect Bayesian equilibria exist and determine a unique equilibrium wage for the worker, profit for the inside firm, and level of total surplus.

To characterize how total surplus and its division into wage and profit depend on the structure of information, it is helpful to study two related quantities: adverse selection and the efficiency contribution. We define adverse selection quantitatively as a real-valued random variable. It is the amount of bad news about general productivity contained in the event that the worker is not retained by the current employer. That is, it is the difference between the expected (general) productivity of the worker conditional on \( T \) alone and conditional on both \( T \) and the event that she switches jobs. Armed with this definition and given the competitive wage setting process, there is another way to think about the quantity of adverse selection: it is the amount that wages are depressed by asymmetric information. In this sense, it is the difference between the equilibrium wage (where outside firms’ wage offers take into account the bad news that the current employer is willing to release the worker at that wage), and the wage they would be prepared to offer if the only relevant information were \( T \). Expected adverse selection therefore quantifies the impact of asymmetric information on the worker’s expected wage.

Trade, that is the worker switching employer, occurs in our model if the inside firm’s estimate of the worker’s productivity in her current employment given \( Q \) is
less than the equilibrium wage. In contrast to Akerlof (1970), trade is not always warranted on efficiency grounds; to maximize total surplus, the worker should remain with her current employer if and only if there is a good match. We call the conditional expectation (given public information) of the contribution to total surplus above the no-trade benchmark the efficiency contribution. The expected efficiency contribution is the difference between total surplus in the market and total surplus obtained if no trade takes place.

In Sect. 4, we characterise the joint distribution of adverse selection, the probability of trade, the efficiency contribution, the inside firm’s profit, and the worker’s wage. The first four are shown to depend on the realization of information \((Q, T)\) only through a scalar statistic that we term apparent match quality. Their joint distribution is governed by three parameters. The wage on the other hand depends on the realization of information \((Q, T)\) through two scalar statistics, and its distribution is governed by five parameters.

To aid interpretation of these characterization results, we discuss a special case of our model where the worker’s outside productivity is linearly related to her inside productivity. In this scalar types case, there are just two free parameters, corresponding to: (a) the quality of public information; and (b) the information gap (Levin 2001), i.e. how much extra information the inside firm has over and above public information. Our model so restricted is effectively a special case of Levin’s model in which values are linearly related.1 We follow Levin in identifying the information gap as a key parameter, and extend his framework by allowing for a non-singular distribution of types. It is straightforward to see how these parameters and (hence) the joint distribution of outcomes change with the information structure. The following two examples illustrate this. Increasing inside information holding public information fixed increases only the information gap. This increases the distribution of adverse selection and inside firm profit by first-order stochastic dominance, and decreases the distribution of the wage by first-order stochastic dominance. On the other hand, increasing the information of both the inside and outside firms (so that the information gap stays constant) increases only the quality of public information. The mean effects are unambiguous: the worker’s expected wage decreases but both the inside firm’s expected profit and expected total surplus increase.

The remainder of the paper presents three applications of this framework. A central question in the literature on the Roy Model is whether there is positive or negative selection into the different sectors—do the best fishermen fish and the best hunters hunt (Roy 1951)? The richness of our model enables us to speak to this issue. In Sect. 5.1, we establish sufficient conditions for there to be positive (negative) selection into the outside market. To aid interpretation, we discuss three cases. In the scalar types case, there can only be negative selection. With non-singular types and an information structure where \(T\) contains the inside firm’s estimate of outside productivity, we obtain the classic Roy model. There can be positive or negative selection into the outside market [according to a single condition familiar from that literature, see e.g. Borjas (1987)] but there is never adverse selection as we define it. However, under a public

---

1 Levin (2001) focuses on the classic durable goods setting and investigates how the structure of information impacts on efficiency. He follows Akerlof (1970) in making a scalar type assumption and investigates conditions under which greater information asymmetries reduce the gains from trade.
information structure where adverse selection falls primarily on workers who should be retained, there is the possibility of an effect reversal (Yule 1903; Simpson 1951); that is, adverse selection at any realisation of public information but, overall, positive selection into the outside market. This possibility has implications for empirical work based on the popular Gibbons and Katz (1991) test for adverse selection.

It might seem plausible (indeed even obvious) that, in the presence of adverse selection, the average wage of released workers will be lower than the average wage of retained workers. But when the type distribution is non-singular this need not be the case. In Sect. 5.2, we establish sufficient conditions for released workers to obtain a premium (penalty) over retained workers. These conditions are related, but not identical, to the conditions from Sect. 5.1; released workers earn a premium because adverse selection falls primarily on workers who are retained.

In Sect. 6, we use our characterization results to address a question familiar from the literature on information design: what ex ante divisions of surplus between the players in the basic game can be achieved as information varies over all possible structures? In our basic game, expected surplus is divided into an expected profit for the inside firm and an expected wage for the worker. As a benchmark, we show (via equilibrium selection) that all feasible expected payoff pairs can be achieved with symmetric information. However, although only one point on the efficiency frontier can be reached, any point in the interior of the feasible set can be achieved as a unique PBE payoff pair in a game with asymmetric information, if there is a non-singular distribution of types. Starting from the first-best point achieved under complete information, it is therefore possible to find public information structures that drive the expected wage down and expected profit up with arbitrarily little efficiency loss. Intuitively, there are information structures that (almost) reveal match value and, in doing so, ensure that (essentially) all adverse selection falls on workers who should be retained. We conclude the discussion by comparing our model and analysis to the information design approach of Bergemann et al. (2015).

The empirical literature on insurance has highlighted some striking findings, e.g. an absence of adverse selection despite there being private information, and the existence of negative adverse selection—sometimes called advantageous selection—where low risks obtain more coverage (Chiappori and Salanie 2000; Finkelstein and McGarry 2006). To date, the theoretical literature has pointed to risk preferences as the primary suspect (de Meza and Webb 2001; Fang and Wu 2018). In Sect. 7, we re-interpret our model in an insurance context to speak to this issue. Specifically, we assume that a consumer with exponential (constant absolute risk aversion) utility seeks insurance for a risky loss, and further that this loss and the consumer’s coefficient of risk aversion are jointly-normally distributed with her inside information $Q$. The characterization results in Sect. 4 apply and have an immediate corollary: advantageous selection occurs if and only if the risky loss and coefficient of risk aversion are sufficiently negatively correlated compared to the amount of asymmetric information. It follows from this condition that scaling up the risk, holding everything else in the model constant, could replace a situation of adverse selection with one of advantageous selection. We conclude by comparing this result to those in Fang and Wu (2018).
2 Model

2.1 Description

We present an employer learning model of the labour market in which firms compete to hire a worker. Specifically, the economy consists of one worker, her current employer (the inside firm) and at least two identical outside firms. The value of the worker to the inside firm, her retained productivity, is denoted by the random variable $V_R$. We will use the words value and productivity interchangeably. The value of the worker to an outside firm, her general productivity, is denoted $V_G$. The match value is $V_M = V_R - V_G$. The joint distribution of $V = (V_G, V_R)$ is exogenously given; it is part of the basic game, in the terminology of Bergemann and Morris (2016). Instead, our focus of interest is the information structure, which specifies a $n_Q + n_T$ dimensional vector $(Q, T)$ of real-valued random variables jointly distributed with $(V_G, V_R)$. Here, $Q$ should be read as information private to the inside firm, $T$ represents public information available to outside firms.

The timing of the game is as follows. Outside firms observe $T$ and then make simultaneous wage offers to the worker. The inside firm, observing $Q$ and the wage offers of the outside firms responds with a wage counter-offer to the worker. The worker accepts one of the wage offers, either from an outside firm or the inside firm. Note that since we do not allow the worker the opportunity to refuse both offers, wages can be negative. Production takes place, the accepted wage offer is paid and payoffs are realized. The worker’s payoff is the wage received, while a firm’s payoff is its profit, i.e. productivity less wages.

A Perfect Bayesian Equilibrium (PBE) of the game requires that the worker, and each firm, follows a sequentially optimal strategy for some Bayes consistent belief. For the worker, her choice of employer must select one of the highest wage offers; for the inside firm, its wage offer must maximise profit given the strategy of the worker and the wage offers of the outside firms; and for each outside firm, its wage offer at each realization $t$ of $T$ must maximise its expected profit given its beliefs about the wage offers made by other outside firms and the strategy of the inside firm.

2.2 Statistical assumptions and notation

Assumption 1 (Asymmetric Information) $T$ is less informative about $(V_G, V_R)$ than $Q$. Specifically, $V_G$ and $V_R$ are conditionally independent of $T$ given $Q$, i.e. $(V_G, V_R) \independent T \mid Q$.

Assumption 1 makes precise the notion of asymmetric information. Note that this admits information structures such as $Q = (V_G, V_R), T = V_G + \epsilon$, where $\epsilon$ is

---

2 The timing in our model of competition for the worker follows the procedure first set out by Greenwald (1986) and which has subsequently been adopted by, e.g., Gibbons and Katz (1991) and Acemoglu and Pischke (1998). An example of an alternative procedure is the simultaneous bidding game studied by Li (2013). Most of the existing employer learning literature does not consider the impact of different information structures. For a recent summary of this literature, see Waldman (2017).
independent of \((V_G, V_R)\). Hence, \(Q\) is more informative than, but need not necessarily contain, \(T\).

It will be convenient to introduce the following notation for the conditional expectations of values given the inside firm’s information

\[
G = E[V_G \mid Q], \quad R = E[V_R \mid Q], \quad M = E[V_M \mid Q].
\]

The random variable \(G\) represents the inside firm’s estimate of the value of the worker to outside firms (and is reminiscent of general human capital), \(R\) is the inside firm’s estimate of the value of the worker if retained, and \(M\) represents its estimate of the quality of its match with the worker. Note that Assumption 1 implies

\[
E[V_G \mid T] = E[E[V_G \mid T, Q] \mid T] = E[E[V_G \mid Q] \mid T] = E[G \mid T].
\]

The first equality follows by the law of iterated expectations, the second by Assumption 1, and the third by the definition of \(G\). Hence, outside firms in estimating their value for the worker need only do so second-hand, via estimating \(G\).

In the sequel we focus exclusively on Gaussian information structures. Let \((V_R, V_G) \sim N(\mu_R, \mu_G, \Sigma_{VR,VG})\). We assume throughout that \(\mu_R = \mu_G\). Consider the game \(G(\mu, \Sigma)\) with \((V_R, V_G, Q, T) \sim N(\mu, \Sigma)\) with \(\mu, \Sigma\) consistent with the basic game (that is with the joint distribution of \(V_R, V_G\)) and with \(\Sigma\) a \(2 + n_Q + n_T\) dimensional positive definite matrix satisfying Assumption 1.³ Let \(\mathcal{G}\) be the set of all such \(\Sigma\) and let \(\mathcal{G} = \{G(\Sigma) \mid \Sigma \in \mathcal{E}\}\). Hence, \(\mathcal{G}\) is the set of games obtainable from the basic game by choosing different information structures respecting Assumption 1.

Since the dimension of \(\Sigma\) may be arbitrarily large and variable across \(\mathcal{E}\), the parameterization of \(\mathcal{G}\) by \(\mathcal{E}\) is a rather inconvenient one. We therefore aim to establish a more convenient parameter space for \(\mathcal{G}\). This will be done in Sect. 3.2, but in preparation for that exercise we introduce a little more notation.

Let \(S_2\) denote the set of \(2 \times 2\) symmetric matrices. For \(A, B \in S_2\), we write \(A \preceq B\) if \(B - A\) is positive definite. The relation \(\preceq\) is sometimes called the Loewner order. The \(2\)-dimensional null matrix is denoted \(0 \in S_2\), hence \(A \succeq 0\) means \(A\) is positive definite.

We write \(\Sigma_{VR,VG}, \Sigma_{RG}\) and \(\Sigma_{RG\mid T}\) respectively to denote the \(2 \times 2\) unconditional and conditional covariance matrices

\[
\Sigma_{VR,VG} = \begin{pmatrix} \sigma_{VR}^2 & \sigma_{VR,VG} \\ \sigma_{VR,VG} & \sigma_{VG}^2 \end{pmatrix}, \quad \Sigma_{RG} = \begin{pmatrix} \sigma_R^2 & \sigma_{RG} \\ \sigma_{RG} & \sigma_G^2 \end{pmatrix}, \quad \Sigma_{RG\mid T} = \begin{pmatrix} \sigma_R^2 | T & \sigma_{RG\mid T} \\ \sigma_{RG\mid T} & \sigma_G^2 | T \end{pmatrix}.
\]

It will often be convenient to refer to Gaussian distributions in terms of regression coefficients. Adopting standard notation, we denote the linear regression coefficients as \((\beta_{GR,T}, \beta_{GT,R})\), so

\[
E[G \mid R = r, T = t] = \text{constant} + \beta_{GR,T} r + \beta_{GT,R} \cdot t.
\]

³ Since \(Q\) and \(T\) may well be linearly related, \((V_R, V_G, Q, T)\) will not generally have a density, however a Gaussian structure can still be represented by its characteristic function \(\phi(u) = \exp \left( u' \mu + \frac{1}{2} u' \Sigma u \right) \), where \(\Sigma\) is a symmetric positive semidefinite matrix. Without loss of generality, we may normalize the means of \(Q\) and \(T\) to be zero.
The regression coefficient $\beta_{GR,T}$ gives the expected change in $G$ for a unit increase in $r$ holding public information $T$ fixed, and plays a central role in our analysis. The covariance parameter $\sigma_{R|T}$ also plays a key role. By the law of total variance, $\sigma_{R|T}^2 = \sigma_{E[V_R|Q]}^2 - \sigma_{E[V_R|T]}^2$. Hence, $\sigma_{R|T}$ can be interpreted as a measure of the difference in the amount of information about the retained value of the worker that is held by the inside and outside firms. We may also write $\beta_{GR,T}^2 \sigma_{R|T}^2 = \sigma_{E[V_G|R,T]}^2 - \sigma_{E[V_G|T]}^2$. Therefore, given $\sigma_{R|T}^2$, we can interpret $\beta_{GR,T}$ as a measure of how much extra information $(R, T)$ contains about $G$ than just the public information $T$. Finally, we reserve `$\epsilon$' to denote “noise” $\epsilon \sim N(0, 1)$ independently distributed from $(Q, T)$. The standard normal PDF and CDF are denoted respectively by $f$ and $F$, and $h : \mathbb{R} \to \mathbb{R}_+$ is the inverse Mills ratio (i.e. the hazard function, $h = \frac{f}{1-F}$).

2.3 The scalar types case

At times, it will be helpful to focus on a scalar types case of the model in which $\Sigma V_R V_G$ is singular. Specifically,

$$V_G = a + b V_R, \quad \text{with } 0 < b < 1.$$ 

We make the natural assumption that the worker’s retained productivity $V_R$ and her general productivity $V_G$ are positively linearly related. Here, $\beta_{GR,T} = b$ independently of the information structure, hence $\beta_{GR,T}$ must always be positive. It is worth emphasising that when $\Sigma V_R V_G$ is non-singular, $V_G$ and $V_R$ can be positively correlated yet $\beta_{GR,T}$ can be negative (for instance, if $T = G + R$, then $\beta_{GR,T} = -1$).

3 Equilibrium

3.1 Equilibrium wage

We start with outside firms’ willingness to pay for a released worker. Under Assumption 1, we can write the willingness to pay of outside firms, i.e. their expected value of the worker conditional on public information $T = t$ and the event that the inside firm is unwilling to retain the worker at wage $w(t) \in \mathbb{R}$, as

$$w^*(t) = E[G \mid T = t, R < w(t)]. \quad (2)$$

This willingness to pay defines a map $w \to w^*$ which under conditions outlined below is a contraction and consequently has a unique solution. This observation lies behind our main equilibrium result.

**Proposition 1** 1. (a) Suppose $\sigma_{R|T} > 0$ and $\beta_{GR,T} < 1$. Perfect Bayesian equilibria exist. In any of these PBE, the equilibrium wage at the realization $T = t$ of public information is the unique solution to the wage equation

$$w(t) = E[G \mid T = t, R < w(t)]. \quad (3)$$
More explicitly, this calculates as

\[ w(t) = E[G \mid T = t] - \beta_{GR,T} \sigma_{R|T} h \left( \frac{E[R \mid T = t] - w(t)}{\sigma_{R|T}} \right). \tag{4} \]

(b) Suppose \( \sigma_{R|T} = 0 \). There exist PBE in which the equilibrium wage is any function satisfying the following conditions: if \( E[G \mid T = t] > E[R|T = t] \), then \( w(t) = E[G \mid T = t] \), and if \( E[G \mid T = t] \leq E[R|T = t] \), then

\[ w(t) \in (-\infty, E[R|T = t]). \]

2. If \( \sigma_{R|T} > 0 \) and \( \beta_{GR,T} \geq 1 \), there is no equilibrium.

**Proof** See Appendix.

When case 1(a) obtains, we will refer to \( t \rightarrow w(t) \) as the equilibrium wage schedule, and use \( W \) to denote the random variable \( W = w(T) \). This equilibrium is familiar and intuitive: there is Bertrand competition and so outside firms make wage offers that, if successful, result in zero expected profit. As always with Bertrand competition, there are many PBE. In particular, there are equilibria where some outside firms make offers that are irrelevant; the number and level of such irrelevant offers is more or less ad hoc devices are added to generate equilibria which exist and are interior to the support of the productivity distribution. See, e.g., Greenwald (1986) who assumes that there is an exogenous separation probability independent of worker productivity.

Part 2 of Proposition 1 states that there is no equilibrium when \( \beta_{GR,T} \geq 1 \). To see why, suppose \( T = M \) (i.e. \( \beta_{GR,T} = 1 \)). Consider a case where the worker is a good match for the inside firm, \( T = t = m > 0 \). An outside firm’s willingness to pay \( E[G \mid T = m, R < \omega] = E[G \mid T = m, G < \omega - m] \) must be less than \( \omega \), and so the wage equation cannot hold. Intuitively, if the inside firm is willing to release a worker who is known to be a good match at wage \( \omega \), then such a worker must be worth less than \( \omega \) to an outside firm.

Case 1(a) of Proposition 1 pins down total surplus, and worker and firm payoffs.\(^5\) Given the equilibrium wage, the total surplus generated is \( TS = 1_{R < W} G + 1_{R \geq W} R \), where \( 1_{R < W} \) denotes the indicator function for the trading event. The worker’s payoff is simply the wage \( W \). Bertrand competition ensures that each outside firm expects to

---

\(^4\) The wage equation in (3) is familiar, for instance, as equation 13.B.6 of Mas-Colell et al. (1995) albeit in a slightly different context and absent public information and non-singular types. Surprisingly this wage equation has not featured in the literature on adverse selection in labor markets. In that literature, more or less ad hoc devices are added to generate equilibria which exist and are interior to the support of the productivity distribution. See, e.g., Greenwald (1986) who assumes that there is an exogenous separation probability independent of worker productivity.

\(^5\) In case 1(b) of Proposition 1, although there are multiple equilibria with wages varying across the equilibria, the allocation is the same in all equilibria.
make zero profit. The inside firm expects a profit of \( R - W \) if it retains the worker and zero otherwise, which we write as \((R - W)^+\). Hence, total surplus is split between the worker and inside firm as \( TS = W + (R - W)^+ \). In Sect. 4, we show that expected values of these quantities depend on the information structure only through *adverse selection* and the *efficiency contribution*. We define these, and other useful quantities, below.

**Definition 1** Suppose, \( \sigma_{R|T} > 0 \) and \( \beta_{GR,T} < 1 \).

1. The *probability of trade* is defined as
   \[
   PR = \Pr[R < W \mid T].
   \] (5)

2. The quantity of *adverse selection* is
   \[
   AS = E[G \mid T] - W.
   \] (6)

3. The *efficiency contribution* (relative to the no-trade status quo) is defined as
   \[
   EC = E[(G - R)1_{\{R < W\}} \mid T] = E[(-M)1_{\{R < W\}} \mid T].
   \] (7)

4. The *inside firm profit* is defined as
   \[
   \Pi = E[(R - W)^+ \mid T].
   \] (8)

The probability of trade defined in Part 1 is simply the probability that trade takes place given public information. We can write it alternatively as \( E[1_{\{R < W\}} \mid T] \).

As we define it in Part 2, the quantity of adverse selection is a wage effect. It is the difference between the equilibrium wage which would obtain in a counterfactual game without private information, \( E[G|T] \), and the equilibrium wage in the game itself, \( W \). Given the wage equation, there is another way of expressing it, which justifies the terminology. It is the difference between the average quality (outside productivity) of the population at large, \( E[G|T] \), and the average quality of that part of the population which selects out of the current employment and into an outside firm, \( E[G|T, R < W] \). We will say that there is positive adverse selection if \( E[G|T] > E[G|T, R < W] \) and negative adverse selection if \( E[G|T] < E[G|T, R < W] \).6

Our definition of adverse selection relates to the widely-used Chiappori and Salanie (2000) positive correlation test. Chiappori and Salanie propose that, if there is adverse selection in a competitive insurance market, then one should observe a positive correlation of risk and coverage conditional on publicly available information. In our unit-demand setting, risk corresponds to \(-V_G\), while coverage corresponds to the trading event. Hence, if there is adverse selection, one should observe a negative correlation between the productivity of the worker and the event that the worker

6 In Sect. 5, we will be interested in selection into the aggregate outside market, no longer conditioning on \( T \). There we will say that there is positive selection into the outside market if \( E[G] < E[G|R < W] \) and negative selection into the outside market if \( E[G] > E[G|R < W] \).
changes jobs conditional on publicly available information. That is, our analogue of the Chiappori-Salanie test is $-\text{Cov}\left(1\{R < w(t)\}, V_G \mid T\right) > 0$. Note that we can write $W = \frac{E[G \times 1\{R < W\} \mid T]}{E[1\{R < W\} \mid T]} = \frac{E[G \times 1\{R < W\} \mid T]}{PR}$. It follows that, on multiplying (6) by $PR$, we have

$$AS \times PR = E[G \mid T]E[1\{R < W\} \mid T] - E[G \times 1\{R < W\} \mid T] = -\text{Cov}\left(1\{R < w(t)\}, V_G \mid T\right).$$

Our measure of the quantity of adverse selection therefore has the same sign as the Chiappori-Salanie test, or what one might term the Chiappori-Salanie covariance of trade and quality.  

In Part 3, we define an expected efficiency measure relative to the no-trade status quo given public information. If the worker remains at the inside firm, the surplus that is generated is $E[R \mid T]$. Hence, what we call the efficiency contribution, is the additional social surplus achieved in the event that worker is released to an outside firm.

**Remark 1** Suppose $\sigma_{R \mid T} > 0$ and $\beta_{G,R,T} < 1$, then

$$EC = E[(W - R)^+ \mid T] \geq 0.$$

**Proof** See Appendix.

Remark 1 can be understood in the original Akerlof context of the sale of a second-hand car, where $W$ is the sale price and $R$ is the owner’s valuation. $EC$ becomes the increase in surplus obtained by the seller through access to the second-hand market. What we define as inside firm profit in Part 4 will be shown to depend on the information structure only through this (positive) efficiency contribution, and adverse selection.

### 3.2 Parameterization of information structures

We conclude this section by using Proposition 1 to reduce the parameter space of all Gaussian information structures to a more convenient representation.

Equation (4) establishes that the value taken by the wage at $T = t$ depends on the information structure only through the conditional distribution of $(G, R)$ given $T = t$. This is intuitive: this conditional distribution is the basis upon which outside firms estimate their valuations for the worker in the event that the worker is released by the inside firm. Given our assumption of Gaussian information structures, the conditional distribution of $(G, R)$ given $T = t$ is parameterised by the conditional means $(E[G \mid T = t], E[R \mid T = t])$ which depend on the realisation $t$, and the conditional variance matrix $\Sigma_{RG \mid T}$ which, of course, in our Gaussian framework does not depend on the realisation $t$. The import of this observation is that the wage

7 Of course, the Chiappori-Salanie test is usually termed the positive correlation test. Since given $T$ the conditional variance of the trading event $1\{R < w(t)\}$ is $\text{var}(1\{R < W\} \mid T) = (E[1\{R < W\} \mid T] - (E[1\{R < W\} \mid T])^2 = PR - PR^2$, the conditional correlation between the trading event $1\{R < W\}$ and $-V_G$ can be written as $\frac{\text{AS} \times PR}{\sqrt{V_G \mid T \cdot PR(1 - PR)}}$. 

Springer
$W = w(T)$ depends on the realization of $T$ (only) through the two dimensional statistic $(\mathbb{E}[R|T], \mathbb{E}[G|T])$ while the functional form of this mapping depends (only) on the conditional covariance matrix $\Sigma_{RG|T}$. Hence, the distribution of the wage is parameterized by $(\Sigma_{RG|T}, \Sigma_{E[R|T]E[G|T]})$. Evidently, the trading event and all other outcomes of interest are determined by the joint distribution of $(W, R, G)$, so the joint distribution of $(R, G)$ parameterized by $\Sigma_{RG}$ will also be relevant. The law of total variance applied to Gaussian distributions may be written

$$\Sigma_{RG} = \Sigma_{RG|T} + \Sigma_{E[R|T]E[G|T]}.$$ 

Therefore specifying any two of $\Sigma_{RG}$, $\Sigma_{RG|T}$, $\Sigma_{E[R|T]E[G|T]}$, determines the third. We conclude and summarise this discussion as follows.

**Proposition 2** (Information Parameter Space) Given the basic game $(V_R, V_G) \sim \mathcal{N}(\mu_{V_R}, \mu_{V_G}, \Sigma_{V_R}V_G)$, the set of games with Gaussian information structures satisfying Assumption 1 can be represented by the parameter space

$$\mathcal{P}_{\Sigma_{V_R}V_G} = \{(\Sigma_{RG}, \Sigma_{RG|T}) \in \mathcal{S}_2 \times \mathcal{S}_2 : 0 \leq \Sigma_{RG|T} \leq \Sigma_{RG} \leq \Sigma_{V_R}V_G\}. \quad (9)$$

**Proof** See proceeding discussion. \qed

Given the law of total variance, the set of information structures may also be represented by

$$\mathcal{P}^+_{\Sigma_{V_R}V_G} = \{(\Sigma_{E[R|T]E[G|T]}, \Sigma_{RG}) \in \mathcal{S}_2 \times \mathcal{S}_2 : 0 \leq \Sigma_{E[R|T]E[G|T]} \leq \Sigma_{RG} \leq \Sigma_{V_R}V_G\}. \quad \mathcal{P}^+_{\Sigma_{V_R}V_G}$$

has the following interpretation. To belong to the parameter space, the information structure must have these properties: the informed party has no better than perfect information about the values ($\Sigma_{RG} \leq \Sigma_{V_R}V_G$); the uninformed party has worse information than the informed party ($\Sigma_{E[R|T]E[G|T]} \leq \Sigma_{RG}$); and there is weakly better than no public information ($\mathcal{T} \leq \Sigma_{E[R|T]E[G|T]}$). To understand this, it is perhaps helpful to remark that, since the means of $(G, R)$ are the same as the means of $(V_G, V_R)$, $\Sigma_{RG} \leq \Sigma_{V_R}V_G$ is equivalent to $E[(aG + bR)^2] \leq E[(aV_G + bV_R)^2]$. Since $aG + bR = E[aV_G + bV_R | Q]$, this inequality follows in the familiar way from an application of Jensen’s inequality.

An important part of our narrative in Sect. 6 will revolve around a distinction between information structures with a zero information gap, $\sigma_{R|T} = 0$, and information structures with a strict information gap, $\sigma_{R|T} > 0$. To this end, we introduce the notation $\mathcal{P}^+_{\Sigma_{V_R}V_G} \subset \mathcal{P}_{\Sigma_{V_R}V_G}$ to denote information structures with a strict information gap (and for which a unique equilibrium wage exists)

$$\mathcal{P}^+_{\Sigma_{V_R}V_G} = \{(\Sigma_{RG}, \Sigma_{RG|T}) \in \mathcal{P}_{\Sigma_{V_R}V_G} : \sigma_{R|T} > 0, \beta_{GR,T} < 1\}. \quad \mathcal{P}^+_{\Sigma_{V_R}V_G}$$

We conclude this section with a final remark.
Remark 2 Proposition 2 implies that any information structure \((Q, T)\) can be represented equivalently by replacing \(Q\) with \((G, R)\), i.e. \(G = \mathbb{E}[V_G | Q]\) and \(R = \mathbb{E}[V_R | Q]\), and \(T\) by

\[
(T_1, T_2) = (a_1 G + b_1 R + c_1 \varepsilon_1, a_2 G + b_2 R + c_2 \varepsilon_2),
\]

where \(\varepsilon_1\) and \(\varepsilon_2\) are independent and identically distributed standard normal noise terms that are also distributed independently of \((G, R)\).

Proof See Appendix.

\[\square\]

4 Characterizing the joint distribution of outcomes

We have seen from Proposition 1 that the parameters \(\sigma_{R|T}\) and \(\beta_{GR,T}\) are of paramount importance in determining whether equilibria exist and whether they are unique. In this section we show that, together with one other parameter, \(\sigma_{E[M|T]}\), they also determine the joint distribution of adverse selection, probability of trade, efficiency contribution and inside firm profit as defined in Definition 1.

4.1 Representation and comparative statics

Our first result highlights the important role played by \(E[M|T]\)—what we term the apparent match quality.

Proposition 3 For information structures in \(\mathcal{P}_{\Sigma_{VRV_G}}^+\), the outcomes \((AS, PR, EC, \Pi)\) depend on the realization of public information only through the scalar statistic \(E[M|T]\). The distribution of these outcomes has as support an arc in \(\mathbb{R}^4\). Specifically, there exists a map \((as, pr, ec, \pi) : \mathbb{R} \to \mathbb{R}^4\) such that

\[
AS = as(E[M|T]), \quad PR = pr(E[M|T]), \quad EC = ec(E[M|T]), \quad \Pi = \pi(E[M|T]).
\]

Furthermore, as is implicitly defined by

\[
as(x) = \beta_{GR,T} \sigma_{R|T} h \left( \frac{as(x) + x}{\sigma_{R|T}} \right),
\]

and

\[
pr(x) = 1 - F \left( \frac{as(x) + x}{\sigma_{R|T}} \right),
\]

\[
ec(x) = \sigma_{R|T} \Psi \left( \frac{x + as(x)}{\sigma_{R|T}} \right),
\]

\[
\pi(x) = \sigma_{R|T} \Psi \left( \frac{x - as(x)}{\sigma_{R|T}} \right) = x + as(x) + ec(x),
\]
where $\Psi(x) = E[(\varepsilon - x)\varepsilon]$. If $0 \leq \beta_{GR,T} < 1$ (positive adverse selection), then $as$ and $\pi$ are increasing, while $pr$ and $ec$ are decreasing. Hence, $AS, -PR, -EC, \Pi$ and $E[M|T]$ are all comonotone. If $\beta_{GR,T} < 0$ (negative adverse selection), then $AS, PR, EC, \Pi$ and $-E[M|T]$ are all comonotone.

**Proof** See Appendix.

Proposition 3 establishes that adverse selection, the probability of trade, the efficiency contribution, and inside firm profit depend on the realization of information $(Q, T)$ only through the scalar statistic $E[M|T]$. The magnitude of adverse selection $|AS|$ and inside firm profit increase in this statistic, while the probability of trade and the efficiency contribution decrease. Intuitively, there is a genuine reason for trade if $E[M|T]$ is small (negative) but not if it is large (positive).

The vector-valued function $(as, pr, ec, \pi)$ is parameterised by a two-dimensional subset of $P^+_{\Sigma V_R V_G}$. It can be represented by the parameters $\sigma_{R|T}, \beta_{GR,T} = \frac{\sigma_{GR,T}}{\sigma_{R|T}}$. Since the mean of $E[M|T]$ is independent of the information structure, its distribution is determined by the single parameter $\sigma_{E[M|T]}$. Hence, recalling that $\varepsilon \sim N(0, 1)$, we have

$$(AS, PR, EC, \Pi, \varepsilon) \overset{d}{=} (as(\sigma_{E[M|T]} \varepsilon), pr(\sigma_{E[M|T]} \varepsilon), ec(\sigma_{E[M|T]} \varepsilon), \pi(\sigma_{E[M|T]} \varepsilon)),$$

where $\overset{d}{=}$ denotes equal in distribution. Our next result considers comparative statics with respect to the three parameters $\sigma_{R|T}, \beta_{GR,T}$ and $\sigma_{E[M|T]}$ when $(AS, PR, EC, \Pi)$ is viewed as a function of $\varepsilon$ as in the representation (14).\(^8\)

**Proposition 4** Consider information structures in $P^+_{\Sigma V_R V_G}$.

1. For $AS$ and $\Pi$:
   
   (a) Increasing either $\sigma_{R|T}$ or $\beta_{GR,T}$ increases the distribution by first-order stochastic dominance. Hence, $E[AS]$ and $E[\Pi]$ increase.
   
   (b) Increasing $\sigma_{E[M|T]}$ produces a single-crossing from below at $\varepsilon = 0$. With larger values of $\sigma_{R|T}$, $AS$ and $\Pi$ are larger when $\varepsilon$ is large and smaller when $\varepsilon$ is small. Both $E[AS]$ and $E[\Pi]$ increase.

2. For $PR$ and $EC$:
   
   (a) Increasing $\sigma_{R|T}$ produces a single-crossing in $PR$. With larger values of $\sigma_{R|T}$, $PR$ is larger when $\varepsilon$ is large and smaller when $\varepsilon$ is small.
   
   (b) Increasing $\beta_{GR,T}$ decreases the distribution by first-order stochastic dominance. Hence, $E[PR]$ and $E[EC]$ decrease.
   
   (c) Increasing $\sigma_{E[M|T]}$ produces a single-crossing at $\varepsilon = 0$. With larger values of $\sigma_{E[M|T]}$, $PR$ and $EC$ are larger when $\varepsilon < 0$ and smaller when $\varepsilon > 0$. $E[EC]$ increases as $\sigma_{E[M|T]}$ increases.

\(^8\) An alternative would to be represent $(AS, PR, EC, \Pi)$ as a function of a uniformly distributed random variable, rather than unit Normal. This would correspond to plotting the functions $(AS, PR, EC, \Pi)$ in terms of their quantiles.
Proposition 4 is (partly) illustrated in Fig. 1, where we use the representation in (14) to plot each of $(AS, PR, EC, \Pi)$ against $\varepsilon$. The lighter-shaded grey plots are all drawn for the same set of parameters: $\sigma_{R|T} = 1$, $\beta_{GR,T} = 0.5$, and $\sigma_{E[M|T]} = 1$. The darker-shaded plots illustrate the distributional consequences of lower parameter values: in the black plots $\sigma_{R|T} = 0.4$, in the red plots $\beta_{GR,T} = 0.3$, and in the blue plots $\sigma_{E[M|T]} = 0.5$. The figure illustrates that, for each outcome and each parameter
change, there is generally a shift according to first-order stochastic dominance, or a single crossing.

Panels (a) and (b) focus on adverse selection and profit. An increase in either $\sigma_{R|T}$ or $\beta_{GR.T}$ increases adverse selection and profit at all values of $\varepsilon$. An increase in $\sigma_{E[M|T]}$ increases adverse selection and profit for $\varepsilon > 0$ (positive apparent match quality) but decreases them for $\varepsilon < 0$ (negative apparent match quality). Despite this single crossing, Proposition 4 Part 1(a) establishes that $E[AS]$ and $E[\Pi_1]$ increase. Panel (c) and (d) focus on the probability of trade and efficiency contribution. An increase in $\beta_{GR.T}$ decreases the probability of trade and efficiency contribution at all values of $\varepsilon$. An increase in $\sigma_{E[M|T]}$ decreases the probability of trade and efficiency contribution for positive apparent match quality but increases them for negative apparent match quality. Despite this single crossing, Proposition 4 Part 2(c) establishes that $E[EC]$ increases. The effect of $\sigma_{R|T}$ on the probability of trade and efficiency contribution is somewhat more complex. There is a single crossing from below but we have no general result on the impact on $E[PR]$ or $E[EC]$.

We have established that the joint distribution of $(AS, PR, EC, \Pi)$ is determined by $(\sigma_{R|T}, \beta_{GR.T}, \sigma_{E[M|T]})$, and considered comparative statics with respect to each parameter. To complete this discussion, we collect results for the equilibrium wage.

**Corollary 1** Consider information structures in $P_{\Sigma_{V_R V_G}}^{+}$. The equilibrium wage $W$ (and hence total surplus $TS = W + \Pi$) depends on the realization of public information through two scalar statistics: apparent match quality $E[M|T]$, and apparent outside productivity $E[G|T]$. Specifically,

$$W = E[G | T] - as(E[M | T]).$$

**Proof** Immediate from Proposition 3 and Definition 1. □

Corollary 1 highlights that the equilibrium wage depends on the realization of information $(Q, T)$ only through two scalar statistics $E[M|T]$, and $E[G|T]$. The distribution of $W$ is determined by five parameters: the three parameters that determine the distribution of $AS$, $(\sigma_{R|T}, \beta_{GR.T}, \sigma_{E[M|T]})$, plus a measure of how much information is available publicly about $V_G$, $\sigma_{E[G|T]}$, and a covariance term, $\sigma_{E[G|T]E[M|T]}$.

Our next result uses the representation in (15) to establish bounds on the equilibrium wage.

**Proposition 5** Consider information structures in $P_{\Sigma_{V_R V_G}}^{+}$. For $\beta_{GR.T} \geq 0$, the equilibrium wage satisfies the bounds

$$W_0 - \sigma_{R|T} \frac{\beta_{GR.T}}{1-\beta_{GR.T}} \sqrt{\frac{2}{\pi}} \leq W \leq W_0,$$

where $W_0 = E[G|T] - \frac{\beta_{GR.T}}{1-\beta_{GR.T}} (E[M|T])^+$. For $\beta_{GR.T} \leq 0$, the inequalities are reversed.

**Proof** See Appendix. □
Evidently, when $\beta_{GR,T} = 0$ the bounds in Proposition 5 are tight: $W = E[G | T]$ and $E[W] = E[G]$. More interestingly, the Proposition asserts that, when the information gap $\sigma_{R|T}$ becomes small, the wage is approximately equal to $W_0$. We will return to the consequences of this fact in Sect. 6, where we explore the distribution of expected total surplus.

We now turn to comparative statics. We view the distribution of wages to be determined by the parameterization $(\sigma_{R|T}, \beta_{GR,T}, \sigma_{E[M|T]}, \sigma_{E[G|T]}, \sigma_{E[G|T]E[M|T]})$ and document comparative statics in the usual way of identifying effects consequent to the change in one parameter holding the others fixed. Clearly, results for $(\sigma_{R|T}, \beta_{GR,T}, \sigma_{E[M|T]})$ carry over from Proposition 4 with a change of sign. For ease of reference, we state them along with the effect of $\sigma_{E[G|T]}$ and $\sigma_{E[G|T]E[M|T]}$ in the following result.

**Proposition 6** Consider information structures in $\mathcal{P}^{+}_{\Sigma_{VR}V_G}$. For the equilibrium wage $W$:

(a) Increasing either $\sigma_{R|T}$ or $\beta_{GR,T}$ decreases the distribution by first-order stochastic dominance. Hence, $E[W]$ decreases.

(b) Increasing $\sigma_{E[M|T]}$ decreases $E[W]$.

(c) Increasing $\sigma_{E[G|T]}$ increases the variance of $W$. $E[W]$ is unchanged.

(d) Increasing $\sigma_{E[G|T]E[M|T]}$ reduces variance of $W$. $E[W]$ is unchanged.

**Proof** See Appendix.

### 4.2 The scalar types case

In this subsection, we use the scalar types case of the model to show how parameters and the joint distribution of outcomes change as the information structure varies in two obvious ways—what happens if there is: more inside information, holding public information fixed; and more public information, holding the extent of asymmetric information fixed?

Recall that, in the scalar types case, $V_G = a + b V_R$ with $0 < b < 1$. Since $\beta_{GR,T} = b$ independently of the information structure, there are just two free parameters, both of which have natural economic interpretations:

- $\sigma_{R|T}$ is the information gap (c.f. Levin 2001). It is a measure of the difference in the amount of information about the scalar type held by the inside and outside firms.\(^9\)
  Hence, with inside information fixed, $\sigma_{R|T}$ decreases with more public information. With public information fixed, it increases with more inside information.
- $\sigma_{E[M|T]} = (1 - b) \sigma_{E[R|T]} = (1 - b) \sigma_{E[V_R|T]}$ is a measure of how much information is available publicly about the scalar type.\(^10\) It is independent of the amount of inside information.

---

\(^9\) Recall that by the law of total variance, $\sigma^2_{R|T} = \sigma^2_{E[V_R|Q]} - \sigma^2_{E[V_R|T]}$. In the scalar types case, $\sigma_{R|T} = 0$ if and only if there is symmetric information, $T = Q$. This is not true in the non-singular types case. For instance, if $T$ contains $R$ but not $G$ there is no information gap but neither is there symmetric information.

\(^10\) In this scalar types case, both $\sigma_{E[G|T]}$ and $\sigma_{E[G|T]E[M|T]}$ are functions of $\sigma_{E[R|T]}$. 

\(\odot\) Springer
Suppose that there is more inside information about the scalar type, holding public information fixed. The information gap $\sigma_{R|T}$ increases but $\sigma_{E[M|T]}$ is unchanged. Applying Propositions 4 and 6, this increases the distribution of $AS$ and $\Pi$ by first-order stochastic dominance and decreases the distribution of $W$, again by first-order stochastic dominance. More inside information keeping outside information the same increases the inside firm’s expected profit and decreases the worker’s expected wage.

Now suppose that both the inside and outside firms acquire more information about the scalar type, such that the information gap $\sigma_{R|T}$ stays constant. The only change is an increase in $\sigma_{E[M|T]}$. This produces single crossings in the distributions of all economic outcomes. More information for all firms again increases the inside firm’s expected profit and decreases the worker’s expected wage. Expected total surplus increases.

Matters are more complex when the distribution of $(VG, VR)$ is non-singular since information is not just about one thing. More information about $VR$ need not imply more information about $VG$ and $VM$. Changes in the information structure will typically move more than one of the relevant parameters at a time.

5 Information and released versus retained workers

In this section, we use the framework developed in Sect. 4 to study two natural questions that arise in the labour market context. First, under which information structures is the average general productivity of released workers higher (or lower) than the average general productivity of retained workers? In particular, can there be an effect reversal (Yule 1903; Simpson 1951)—that is, adverse selection at any $T$ but, on average, positive selection of general productivity into the outside market? Second, under which information structures is the average wage of released workers higher (or lower) than the average wage of retained workers? In particular, can released workers earn a premium over retained workers even in the presence of adverse selection?

5.1 When is there positive selection into the outside market?

A central question in the literature on the Roy model is whether there is positive or negative selection into the different sectors. Our next result speaks to this question.

Proposition 7 Consider information structures in $P_{GR,T}^+$. Suppose $\beta_{GR,T} \geq 0$.

1. If $\sigma_{E[G|T]}E[M|T] < -\beta_{GR,T}\sigma^2_{R|T}$, then there is positive selection into the outside market, i.e. $E[G \mid R < W] > E[G]$.

2. If $\sigma_{E[G|T]}E[M|T] > -\beta_{GR,T}\sigma^2_{R|T}$, then there is negative selection into the outside market, i.e. $E[G \mid R < W] < E[G]$.

Proof See Appendix.

Part 1 of Proposition 7 establishes a sufficient condition for positive selection into the outside market. Note that if both $\beta_{GR,T}$ and $\sigma_{R|T}$ are strictly positive, then there is an effect reversal: $E[G \mid R < W] > E[G]$ in spite of $E[G \mid R < W, T] < E[G|T]$. Part 2 establishes a sufficient condition for negative selection into the outside market.
To interpret this result, it is helpful to return to the scalar types case in Sect. 2.3. The condition in Part 1 becomes \( b(1-b)\sigma_{G|M}^2 < -b\sigma_{R|T}^2 \) which cannot hold, while Part 2 becomes \( b(1-b)(1+\sigma_{R|T}^2) > 0 \) which must always hold. There can only be negative selection into the outside market, and an effect reversal cannot arise. Next, consider non-singular types with public information \( T = G \). In this, the classic Roy model case, Parts 1 and 2 collapse to the single necessary and sufficient condition for positive selection \( \sigma_{GM} < 0 \) (equivalently \( \beta_{RG} < 1 \)). There can now be positive selection into the outside market, although an effect reversal still cannot arise because \( E[G \mid R < W, T] = E[G \mid T] \). Finally, consider non-singular types with \( T = (G - \beta R, M + \sigma \varepsilon) \) and \( 0 < \beta < 1 \). As \( \sigma \to 0, T \) becomes increasingly informative about \( G \) and \( R \) and is fully informative in the limit. Hence, \( \sigma_{R|T} \to 0 \) and \( \sigma_{E[G|T]E[M|T]} \to \sigma_{GM} \). It follows that, as \( \sigma \to 0 \), Part 1 approaches the single necessary and sufficient condition for positive selection, \( \sigma_{GM} < 0 \), from the Roy Model case. The difference is that an effect reversal can arise because \( E[G \mid R < W, T] < E[G \mid T] \).

This possibility of an effect reversal has implications for empirical work. Gibbons and Katz (1991) famously test for adverse selection by comparing wages for workers who are selectively ‘laid off’ with those who become unemployed due to exogenous plant closures. The idea is that with plant closures there is no stigma, so if one can condition on \( T \), wages should on average be \( E[G \mid T] \). In contrast, for workers who are selectively laid off, wages are depressed by adverse selection to \( W \) and \( E[G \mid T] - AS \). Hence, conditioning on \( T \), in our model the Gibbons-Katz test gives a positive result whenever \( AS > 0 \).

But what if \( T \) is not observed? In terms of our model, it follows from the wage equation (competition) that expected wages for workers who are selectively laid off are \( E[W \mid R < W] = E[G \mid R < W] \). Hence, (absent the econometrician observing \( T \)) the Gibbons-Katz test reduces to testing whether there is negative selection into the (aggregated) outside market, i.e. \( E[G \mid R < W] \leq E[G] \). Using Proposition 7, a strongly sufficient condition for this form of the Gibbons-Katz test to find adverse selection when there is indeed (positive) adverse selection is that \( M \) has a degenerate distribution. This is because when \( \sigma_{E[G|T]E[M|T]} \to 0 \) an effect reversal cannot arise. Hence, if the econometrician can neither observe \( T \) nor verify that this sufficient condition holds, the possibility of an effect reversal cautions against using this form of the Gibbons-Katz test.

5.2 When is the average wage of released workers higher than the average wage of retained workers?

It might seem plausible (indeed even obvious) that, in the presence of adverse selection, released workers will always earn less than retained workers. However, this need not

---

11 Relatedly, in our unit-demand model, the well-known positive correlation test of Chiappori and Salanie (2000) becomes a simple test of whether \( AS > 0 \). As stressed by Chiappori and Salanie, since theory implies different observable types should be offered different contracts, one should control for public information and therefore test whether \( E[G|T] - E[G \mid R < W, T] > 0 \). However, if there is unobserved heterogeneity, i.e. \( T \) is poorly observed by the econometrician, this control is difficult and the natural surrogate is to estimate the quantity of adverse selection in the aggregated market constructed by ignoring \( T \), specifically \( E[G] - E[G \mid R < W] \).
be the case. In our model, competition means that outside firms get what they pay for in the sense that $E[W | R < W] = E[G | R < W]$. Since $E[AS] = E[G] - E[W]$, we may write the difference between expected wages for those who leave to outside firms and the expected wage of the population as the difference between the amount of selection in the aggregate market and the aggregate amount of adverse selection, i.e.,

$$E[W] - E[W | R < W] = E[G] - E[G | R < W] - E[AS].$$  

(16)

This observation gives us the following result.

**Proposition 8** Consider information structures in $\mathcal{P}^+_{\Sigma_{VRV_G}}$. Suppose $\beta_{GR,T} \geq 0$.

1. If $\sigma_{E[G | T]E[M | T]} < 0$, then the average wage of released workers is higher than the average wage of retained workers, i.e. $E[W | R < W] > E[W]$.

2. If $\sigma_{E[G | T]E[M | T]} \geq \frac{\beta_{GR,T}}{1 - \beta_{GR,T}} \sigma_{E[M | T]}^2$, then the average wage of released workers is lower than the average wage of retained workers, i.e. $E[W | R < W] < E[W]$.

**Proof** See Appendix.  

Part 1 of Proposition 8 establishes a sufficient condition for the average wage of released workers to be higher than the average wage of retained workers, while Part 2 establishes a sufficient condition for the reverse. Again, to interpret this result, it is helpful to start with the scalar types case in Sect. 2.3. The condition in Part 1 becomes $b(1 - b)\sigma_{E[R | T]}^2 < 0$ which cannot hold, while Part 2 becomes $b(1 - b)\sigma_{E[R | T]}^2 \geq b(1 - b)\sigma_{E[R | T]}^2$ which obviously must always hold. So in the scalar types case, the standard intuition applies: the average wage of released workers is lower than the average wage of retained workers, with or without adverse selection. In the non-singular types case with $T = G$, Parts 1 and 2 collapse to the same single necessary and sufficient condition from Proposition 7. Now the average wage of released workers can be higher than the wage of released workers (if $\sigma_{GM} > 0$), but of course under this information structure there is no adverse selection. In the non-singular types case with $T = (G - \beta R, M + \sigma \varepsilon)$ and $0 < \beta < 1$, as $\sigma \to 0$ Part 1 approaches the sufficient condition $\sigma_{GM} < 0$. When this is satisfied, released workers earn a premium over retained workers even in the presence of adverse selection.

### 6 Information and expected payoffs

In this section, we use our framework to explore the PBE payoff pairs—an expected wage for the worker and an expected profit level for the inside firm—that are possible as the information structure varies over the parameter space.  

If the worker is allocated efficiently on the basis of full information, then the expected total surplus available is $E[\max(V_G, V_R)]$. Conversely, if the worker is allocated on the basis of no information, the expected total surplus available is $E[V_R] = E[V_G]$. We define the set of expected payoffs which allocate an amount of

---

12 Recall that all outside firms make zero expected profit under any information structure.
expected total surplus between these two bounds, subject to the inside firm’s expected profit being non-negative, as

\[ \Omega_{\Sigma_{VRVG}} = \left\{ (E[\Pi], E[W]) \in \mathbb{R}^2 : E[\Pi] \geq 0, E[V_R] \leq E[\Pi] + E[W] \leq E[\max(V_G, V_R)] \right\}. \]

Since both inside firm profit and the efficiency contribution are non-negative, any PBE expected profit-expected wage payoff pair must lie in the set \( \Omega_{\Sigma_{VRVG}} \) and so \( \Omega_{\Sigma_{VRVG}} \) should be read as the entire space of possibilities. The question, then, is which points in the set \( \Omega_{\Sigma_{VRVG}} \) are PBE payoff pairs?

To answer this question, let \( \mathcal{P} \subseteq \mathcal{P}_{\Sigma_{VRVG}} \) and denote the set of expected profit, expected wage pairs achievable as a PBE for some information set in \( \mathcal{P} \) by \( \mathcal{A}(\mathcal{P}) \). Further, let \( \mathcal{A} \) denote the closure of the set \( \mathcal{A} \subset \mathbb{R}^2 \). Using this notation, we can state the following result.

**Proposition 9**

1. For any \( \Sigma_{VRVG} \in \mathcal{S}_2 \), any point in \( \Omega_{\Sigma_{VRVG}} \) is a PBE expected profit-expected wage payoff pair for some information structure in \( \mathcal{P}_{\Sigma_{VRVG}} \). That is,

\[ \mathcal{A}(\mathcal{P}_{\Sigma_{VRVG}}) = \Omega_{\Sigma_{VRVG}}. \]

2. Suppose \( \Sigma_{VRVG} \succ 0 \). Any allocation in \( \Omega_{\Sigma_{VRVG}} \) can be approached as the limit of a sequence of information structures in \( \mathcal{P}_{\Sigma_{VRVG}}^+ \). That is,

\[ \mathcal{A}(\mathcal{P}_{\Sigma_{VRVG}}^+) = \Omega_{\Sigma_{VRVG}}. \]

For any information structure in \( \mathcal{P}_{\Sigma_{VRVG}}^+ \), \( E[W] \neq E[V_G] \) implies \( E[TS] < E[\max(G, R)] \leq E[\max(V_G, V_R)] \).

3. Suppose \( \Sigma_{VRVG} \) is singular and \( V_G = a + bV_R \), with \( 0 < b < 1 \). For any information structure in \( \mathcal{P}_{\Sigma_{VRVG}}^+ \), the PBE expected profit-expected wage payoff pair satisfies

\[ E[W] \leq E[V_G] - b E[\Pi]. \]

Hence,

\[ \mathcal{A}(\mathcal{P}_{\Sigma_{VRVG}}^+) \neq \Omega_{\Sigma_{VRVG}}. \]

**Proof** See Appendix. \( \square \)

Proposition 9 is illustrated in Fig. 2, where the two shaded regions together constitute the set \( \Omega_{\Sigma_{VRVG}} \). Part 1 of the proposition states that any point in this (dark and light-) shaded region is a PBE expected profit-expected wage payoff pair for some
information structure in $\mathcal{P}_{\Sigma_V R G}$. To see why, note that we allow for no information gap, $\sigma_{R|T} = 0$. In this case, when the inside firm is known to have a valuation $R$ that is higher than the valuation $G = a + b R$ of outside firms, any offer by an outside firm that is less than $R$ will be matched by the inside firm and so will be unsuccessful. Hence, the equilibrium wage can lie anywhere in the interval $(-\infty, R]$, with corresponding equilibrium profit $[0, \infty)$. By varying the quality of information (maintaining the assumption that it is held symmetrically so that $\sigma_{R|T} = 0$) it is therefore possible to achieve any point in the set $\Omega_{\Sigma_V R G}$ through equilibrium selection.

Part 2 of the proposition considers only asymmetric information structures. Here we establish that when $\Sigma_{V R G}$ is non-singular, any point in $\Omega_{\Sigma_{V R G}}$ can be reached as a limit of a PBE expected profit-expected wage payoff pair for some information structure in $\mathcal{P} \Sigma_{V R G}$ with a strict information gap. Only one point on the first-best efficiency frontier can be achieved, namely the PBE where $E[W] = E[V_G]$ which is obtained when $Q = (V_G, R)$ and $T = G$. This is Point A in Fig. 2. Hence, non-zero expected adverse selection implies some, but only infinitesimal, loss of surplus.

To see the intuition, consider an information structure of the form $Q = (V_G, V_R)$ and $T = (G - \beta R, M + \sigma \varepsilon)$, where $\beta < 1$. As $\sigma \to 0$, $T$ becomes increasingly informative about $G$ and $R$ and is fully informative in the limit. Hence, $\sigma_{R|T} \to 0$ and $\sigma_{E[M|T]} \to \sigma_M$. However, $\beta_{GR,T} = \beta$ independently of $\sigma$. It is this feature—the ability to vary $\beta_{GR,T}$ independently of $\sigma$—that means that $E[AS]$ can remain large in absolute value, even as $\sigma \to 0$. At one extreme with $\beta \to -\infty$ and $\sigma \to 0$, $E[AS]$ is negative and we approach the best outcome for the worker and the worst for the firm. This is point B in Fig. 2. At the other extreme with $\beta \to 1$ and $\sigma \to 0$, $E[AS]$ is positive and we approach the best outcome for the firm and the worst for the worker. Here, the direction of travel is shown by the arrow in Fig. 2. In this information

---

13 Note that in illustrating all parts of Proposition 9 on the same diagram we are implicitly assuming that $E[\max(V_G, V_R)]$ is the same in the scalar and non-singular type cases.

14 Recall Remark 2. Here, $a_1 = 1$, $b_1 = -\beta$, $c_1 = 0$ and $a_2 = -1$, $b_2 = 1$, $c_2 = \sigma$. 
structure, \( Q = (V_G, V_R) \). By reducing the quality of the inside firm’s information, it is possible to achieve any point in the interior of the set \( \Omega \Sigma V_R V_G \). Note that, in contrast to Part 1, this is not through equilibrium selection; each point is a unique PBE payoff pair in a game with asymmetric information.

It may initially seem counter-intuitive that large quantities of expected adverse selection can be imposed with only an arbitrarily small efficiency loss. Note that, by Proposition 3, if \( E[AS] > 0 \), then \( AS \) and \( -PR \) are comonotone. Hence, \( \text{Cov}(AS, PR) < 0 \) which implies that \( E[AS | R < W] < E[AS | R \geq W] \)—in other words, adverse selection always falls mainly on the trades that do not take place. Part 2 of Proposition 9 establishes that there are information structures where essentially all adverse selection falls on trades that should not take place. That is, it falls on workers with \( M > 0 \) whom it is efficient to retain.

The final part of Proposition 9 provides a comparison with the scalar types case. Here, \( \Sigma V_R V_G \) is singular and \( V_G = a + bV_R \), with \( b \) fixed and satisfying \( 0 < b < 1 \). The dark-shaded region in Fig. 2 is now excluded from the set of PBE payoff pairs that are possible under information structures in \( P_+^{\Sigma V_R V_G} \). The striking result in Part 2—that any point in the interior of \( \Omega \Sigma V_R V_G \) can be achieved as a unique PBE payoff pair in a game with asymmetric information—no longer holds. To see why, observe from the representation in (10) that, with \( b \) fixed and satisfying \( 0 < b < 1 \), \( AS \) cannot be negative. This immediately rules out the triangle \( ABC \). Starting from no inside information (at point C), expected profit can increase as more information becomes available but only at the expense of reducing the expected wage through adverse selection, the terms of the trade-off being determined by the fixed parameter \( b \). This rules out the triangle \( ACD \).

Given the similarity of some results, we conclude this section by briefly comparing our model and analytical approach to that of Bergemann, Brooks, and Morris (2015, hereafter BBM). In BBM, the monopoly seller’s value is zero and (in the standard interpretation) there is one buyer who receives a signal which determines her value but the seller does not know the realisation of this signal at the time he sets the price. There is asymmetric information but it is the buyer who has the private information not the seller. We can also interpret our framework as a one-seller, one-buyer model. Here, the monopoly seller’s value is a random variable \( R \) and there is one buyer who receives a signal \( T \). The seller knows the realisation of this signal at the time he sets the price (or more accurately makes a take-it-or-leave-it offer). There is asymmetric information but it is the seller who has the private information not the buyer.

BBM study the combinations of consumer and producer surplus that can be achieved by considering all possible information structures. They proceed by characterizing certain convex sets by their extreme points (which, by mixing, therefore represent the entirety of information structures), and show that any feasible distribution of surplus

---

15 Interpreting our framework as a one-seller, one-buyer model is not quite a relabelling exercise. In the labour market setting, it is natural to ask how the information structure affects the split of expected total surplus between the worker and the inside firm. In contrast, in the one-seller, one-buyer setting, the issue of how the surplus is split between the players is not interesting. The take-or-leave-it offer means that the buyer never receives any surplus. It is trivial to see what would be best for the seller: complete information, since that delivers first-best expected total surplus.

16 Since mixtures of normally distributed random variables are not normal, we do not allow mixing.
is possible. We also study the combinations of expected payoffs that can be achieved by considering all possible information structures, and show in Part 1 of Proposition 9 that any feasible distribution of surplus is possible. In contrast to BBM, however, our proof rests on a rather straightforward equilibrium selection argument that can be made exclusively under symmetric information. Given this, in order to analyse the impact of asymmetric information, we take it to be an essential prerequisite to assume a strict information gap. This restriction to a strict information gap, together with the fact that mixtures of Gaussian distributions are not Gaussian means that the familiar mathematical strategy of information design—to characterize a compact convex set using the Krein-Milman theorem via the convex hull of its extreme points—is not available. In Part 2 of Proposition 9 we show that even with a strict information gap, any point in the set of possible surplus divisions can be achieved in the limit if there is a non-singular distribution of types. This result is attained via a proof strategy very different from the convex hull one. We characterise the set of achievable payoff pairs by first characterizing what is attainable in the limit as the information gap becomes vanishingly small. This yields a dense subset of the set desired to be characterised.

7 Advantageous selection in insurance markets

We have framed our model in terms of a labour market but it is straightforward to re-interpret it in the context of an insurance market, where the consumer takes the role of the informed inside firm and uninformed insurance companies take the role of outside firms. We do so here, in order to facilitate comparisons with the wider literature. Specifically, we show how the characterization results in Sect. 4 can contribute to discussions of negative adverse selection, or what the literature has tended to call advantageous selection (c.f. de Meza and Webb 2001; Finkelstein and McGarry 2006; Fang and Wu 2018). We are grateful to a referee for suggesting this application. It will be helpful to use negative adverse selection and advantageous selection interchangeably in this section.

Translated to the competitive insurance market setting, the model runs as follows. Let $A$ denote the Arrow-Pratt coefficient of absolute risk aversion of a consumer facing the risky prospect $y - X$. We interpret $y$ as initial income, assumed to be known, and $X$ as a risky but insurable loss. From Arrow (1971), the certainty equivalent of the risky prospect $y - X$ is approximately equal to

$$R = \mathbb{E}[y - X|Q] - \frac{1}{2} A \sigma_X^2|Q|,$$

where $Q$ denotes the consumer’s private information which we assume contains $A$. The approximation is exact in the case of Gaussian losses and exponential (constant absolute risk aversion) utility. We adopt this CARA Gaussian framework here. A risk-neutral insurer taking on the risky prospect would, if it had the same information $Q$ as the consumer, have a valuation of

$$G = \mathbb{E}[y - X|Q].$$

17 We are grateful to a referee for suggesting this application. It will be helpful to use negative adverse selection and advantageous selection interchangeably in this section.
We maintain Assumption 1 and assume that $A$, $X$, $Q$, $T$ are joint-Normal.

We know from (10) that the sign of $\beta_{GR,T}$ determines the sign of adverse selection. This immediately gives us a necessary and sufficient condition for advantageous selection.

**Corollary 2** There is advantageous selection in the competitive insurance market if and only if $\beta_{GR,T} < 0$, which is equivalent to

$$- \frac{\sigma_{AX|T}}{2} > \frac{\sigma_{X|T}^2}{\sigma_{X|Q}^2}. \quad (17)$$

**Proof** See Appendix. \hfill \qed

To read this, first note that, since $Q$ is more informative about $X$ than is $T$, $\frac{\sigma_{X|T}^2}{\sigma_{X|Q}^2}$ is greater than one. This term on the right hand side of (17) is a measure of how asymmetric is information about the risky loss $X$. Corollary 2 states that advantageous selection occurs if the Arrow-Pratt coefficient of risk aversion $A$ and the risky loss $X$ are sufficiently negatively correlated compared to (this measure of) the amount of asymmetric information.

Two further observations are immediate from Corollary 2. First, there cannot be advantageous selection when the consumer has perfect information about the loss, i.e. $\sigma_{X|Q}^2 = 0$. The reason is intuitive: in this case, differences in risk aversion do not affect behaviour. Since the consumer already knows the extent of the loss $X$ when contracting she values disposing of the risk in exactly the same way irrespective of her degree of risk aversion.

Second, scaling up the risk [i.e. $X \rightarrow \alpha X$ for some $\alpha \in (1, \infty)$] could replace a situation of adverse selection with one of advantageous selection. This is easiest to see if we suppose that public information is null so that $\sigma_{AX|T} = \sigma_{AX}$. Clearly, (17) can only hold if $\sigma_{AX} < 0$, implying that the low risk consumers are also the more risk averse consumers. Suppose therefore that $\sigma_{AX} < 0$ and consider the experiment of doubling all losses, $X$ becomes $2X$, keeping everything else the same. The left hand side of (17) becomes twice as big but the right hand side is unchanged (the left hand side is homogeneous of degree one in $X$ whereas the right hand side is homogeneous of degree zero). Therefore, starting from a small risk and making it more acute, will eventually replace adverse selection with advantageous selection. The simple intuition is that scaling the risk up while holding the coefficient of risk aversion constant increases the absolute value of the risk premium (since the risk premium is the product of the coefficient of risk aversion with (half) the variance of the risk). Eventually, with the scaling, those buying insurance are predominantly those with high risk premia but low risks.

These observations, while consistent with the initial intuition offered in de Meza and Webb (2001) and Finkelstein and McGarry (2006), have a different flavour to more recent theoretical investigations. Proposition 1 of Fang and Wu (2018) asserts that advantageous selection cannot occur in competitive insurance markets unless there is a large load factor (administrative cost to supplying insurance). Evidently,
our presentation above has zero load so Fang and Wu’s proposition is in apparent
contradiction to Corollary 2. They assume that all consumers are risk averse, whereas
our Gaussian assumption evidently admits a (possibly very small) measure of risk
preferring consumers. The assumption of risk aversion is very standard, so one can
hardly object to it. However, our treatment does highlight that there can be a striking
difference between a model in which all consumers are risk averse and one in which all
but an arbitrarily small proportion are. In the first case, it becomes common knowledge
independently of the trading event that the existing position of no-insurance is an
inefficient allocation, whereas in the latter case, it is a matter for inference based on
equilibrium behaviour.

8 Conclusion

This paper has explored how the structure of asymmetric information impacts eco-
nomic outcomes in a natural generalization of Akerlof’s (1970) Lemons model that
allows for a random match component between the good to be traded (a worker) and
the seller (the inside firm), and information structures consisting of both inside and
public information.

We show that the joint distribution of adverse selection, probability of trade, effi-
ciency contribution and inside firm profit has a simple, tractable representation—the
distribution has as support an arc in \( \mathbb{R}^4 \). Specifically, these outcomes depend on infor-
mation \((Q, T)\) only through one scalar statistic that we term apparent match quality.
The distribution of wages depends on two statistics, but the expected wage depends
on information only by its effect on expected adverse selection.

We use this framework to address several applications. First, in the labour market
context we identify conditions under which an effect reversal—adverse selection at
any realisation of public information but, overall, positive selection into the outside
sector—can and cannot arise. In settings where match components may be relevant in
addition to general human capital, this result calls for care in applying and interpreting
empirical tests for adverse selection. Second, our characterization results allow us to
explore the \textit{ex ante} divisions of surplus between players in the basic game that can be
achieved as information varies over all possible structures. In our basic game, expected
surplus is divided into an expected profit for the inside firm and an expected wage for
the worker. We show that, although only one point on the efficiency frontier can be
reached, any point in the interior of the feasible set can be achieved as a unique PBE
payoff pair in a game with asymmetric information, if the distribution of worker types
is non-singular. In a related working paper, we explore how this has implications
for endogenous information structures that arise as strategic choices in a two-period
model. Finally, we reinterpret the model in an insurance context where the analogue
of the general human capital component is the loss to be insured and the matching
component is consumer risk aversion. Our characterization results are straightforward
to apply and bring some simple insights to the recent literature discussing sources of
advantageous selection.
Appendix

It is convenient to collect here some mainly familiar properties of the normal hazard function (inverse Mills ratio).

Lemma A1 (Properties of the normal hazard function) The normal hazard function $h : \mathbb{R} \rightarrow \mathbb{R}_+$, satisfies:

1. Level bounds. $[x]^+ < h(x) < \frac{\sqrt{4+x^2}-x}{2} < |x + \frac{1}{x}|$ on $\mathbb{R}$. This implies
   \[ \lim_{x \to \infty} \frac{h(x)}{x} = \lim_{x \to \infty} h'(x) = 1. \]
2. Convexity bounds. $0 < h' < 1$ on $\mathbb{R}$.
3. Gradient bounds. $h$ is strictly convex and log concave on $\mathbb{R}$. $0 < \frac{d}{dx} \frac{h(x)}{h'(x)} < 1$.
4. Star-shaped. $x \mapsto h(x)$ and $x \mapsto \frac{h(x)}{h'(x)}$ are both strictly star-shaped on $\mathbb{R}$. We call a function $g : \mathbb{R} \rightarrow \mathbb{R}$ with $g(0) > 0$ strictly star-shaped if for all $x \in \mathbb{R}$, $0 < \alpha < 1$, $\alpha g(x) < g(\alpha x)$.

Proof Property 1. The lower bound and larger upper bound are in Gordon (1941). The tighter upper bound was established by Birnbaum (1942).

Property 2. One easily verifies $h'(x) = h(x)(h(x) - x)$ which, given Property 1, is positive. Since convexity implies the derivative is increasing, the final statement of 1. implies $h' < 1$.

Property 3. Convexity was conjectured by Birnbaum (1950), and proved by Sampford (1953). Log-concavity is established in Marshall and Olkin (2007) Lemma B.4 p.437. Marshall and Olkin establish the result using an inequality of Birnbaum. A simpler proof follows from using $h'(x) = h(x)(h(x) - x)$ and using Property 2. That $h/h'$ is increasing follows immediately from the log-concavity of $h$. That $\frac{d}{dx} \frac{h(x)}{h'(x)} < 1$, follows from the convexity of $h$ upon observing that $\frac{d}{dx} \frac{h(x)}{h'(x)} = 1 - \frac{h(x)h''(x)}{(h'(x))^2}$.

Property 4. A continuous positive function is star-shaped on $\mathbb{R}$ if it is star-shaped on $(0, \infty)$. $h$ is star-shaped if $\frac{h'(x)}{h''(x)} \leq \frac{1}{x}$ on $(0, \infty)$. One easily verifies $h'(x) = h(x)(h(x) - x)$. Hence, it suffices to show that $(h(x) - x)x \leq 1$, the result therefore follows from 1. To establish $x \mapsto \frac{h(x)}{h'(x)}$ is star-shaped on $(0, \infty)$, we require that $h/h'$ is decreasing on $x > 0$, this follows from $h$ being star-shaped ($h/x$ is decreasing) and by the convexity of $h$, $\frac{1}{h'}$ is decreasing. Hence, the result follows from the fact that the product of two decreasing positive functions is decreasing. \hfill \square

Material omitted from Sect. 3

Proposition 1.
Proof Part 1(a). First, we establish uniqueness. By the law of iterated expectations, the equilibrium wage equation can be written as

\[ w(t) = E[G \mid T = t, R < w(t)]. \]

Using the linear conditional expectations properties of Gaussian distributions, this becomes, in standard regression notation,

\[ w(t) = E[G \mid T = t] + \beta_{GR,T}E[R - E[R \mid T = t] \mid T = t, R < w(t)]. \]

Since given \( T = t \), \( R - E[R \mid T = t] \) is distributed \( \mathcal{N}(0, \sigma_{R|T}^2) \), and \( E[\varepsilon \mid \varepsilon < w(t)-E[R|T=t]/\sigma_{R|T}] = -f\left(\frac{w(t)-E[R|T=t]}{\sigma_{R|T}}\right) / F\left(\frac{w(t)-E[R|T=t]}{\sigma_{R|T}}\right) \), we obtain

\[ w(t) = E[G \mid T = t] - \beta_{GR,T}\sigma_{R|T}h\left(\frac{E[R \mid T = t] - w(t)}{\sigma_{R|T}}\right). \quad (18) \]

If there were two solutions, \( w(t) > \hat{w}(t) \), to this equation, then by the mean value theorem, for some \( z \) in the interval \( \left(\hat{w}(t), w(t)\right) \), \( w(t) - \hat{w}(t) = \beta_{GR,T}h'\left(\frac{E[R|T=t]-z}{\sigma_{R|T}}\right) \) \( (w - \hat{w}) \), therefore, since \( h' < 1 \), with \( \beta_{GR,T} < 1 \) we have a contradiction. Hence, if there is a solution at \( T = t \), it is unique. That there exists a solution follows from continuity and the fact that with \( \beta_{GR,T} < 1 \), and

\[ \xi(x) = x - E[G \mid T = t] + \beta_{GR,T}\sigma_{R|T}h\left(\frac{E[R \mid T = t] - x}{\sigma_{R|T}}\right), \]

we have \( \lim_{x \to +\infty} \xi(x) = +\infty \) and \( \lim_{x \to -\infty} \xi(x) = -\infty \). These limits are easily seen to follow from the fact that \( x^+ \leq h(x) \leq h(0) + x^+ \).

Part 1(b). If \( \sigma_{R|T} = 0 \), then if \( E[G \mid T] > R \) it is common knowledge that the worker is more valuable to the outside firms than the inside firm. Bertrand competition means the equilibrium wage is the common expected value to outside firms, given their information, i.e. \( E[G \mid R] \). In the event that \( E[G \mid R] \leq R \), outside firms will be unwilling to make wage offers in the interval \( (E[G \mid R], \infty) \) but any other wage offer is optimal since the inside firm will outbid such offers and achieve their maximal payoff of 0.

Part 2. For the game to have an equilibrium the wage equation must have a solution for each possible realisation of public information. For simplicity, take the case \( E[G \mid T = t] = E[R \mid T = t] = z \) then for \( \xi(x) = 0 \) it must be that

\[ \frac{z - x}{\sigma_{R|T}} = \beta_{GR,T}h\left(\frac{z - x}{\sigma_{R|T}}\right) \]

however, since \( h(x) > x \) this is impossible with \( \beta_{GR,T} \geq 1 \). The same argument holds for any case with \( E[G \mid T = t] \leq E[R \mid T = t] \). To complete the proof, note that since \( V_R \) and \( V_G \) are assumed to have the same mean, \( \mu_R = \mu_G \), it is impossible to have \( E[G \mid T = t] > E[R \mid T = t] \) for all \( t \).
Remark 1.

Proof Since,

\[ W = \frac{E[ G \cdot 1_{R < W} | T]}{E[1_{R < W} | T]}, \]

we may write

\[ E[ G \cdot 1_{R < W} | T] = E[W \cdot 1_{R < W} | T], \]

subtracting \( E[R \cdot 1_{R < W}| T] \) from both sides gives

\[ E[(G - R)1_{R < W} | T] = E[(W - R)1_{R < W} | T] \]

\[ = E[(W - R)^+ | T] \geq 0. \]

\[ \square \]

Remark 2.

Proof Since public information impacts on \( \Sigma_1 \) \( E[R|T] \) but not \( \Sigma_{RG} \), any Gaussian signal \( T \) impacts on the distribution of outcomes only through the binary statistic \( (T_1, T_2) = (E[R|T], E[G|T]) \) which is also Gaussian. Hence, there is no loss of generality in reducing the dimensionality of \( T \) to two. It is also without loss of generality to write the first statistic in the form \( T_1 = a_1G + b_1R + c_1\varepsilon_1 \), where \( \varepsilon_1 \sim \mathcal{N}(0, 1) \) independently of \((G, R)\). Similarly, having chosen this representation for \( T_1 \), it is without loss of generality to represent the second statistic as \( T_2 = a_2G + b_2R + c_2\varepsilon_1 + d_2\varepsilon_2 \), where \( \varepsilon_2 \sim \mathcal{N}(0, 1) \) independently of \((G, R, \varepsilon_1)\). With \( c_1, d_2 \neq 0 \), \( (T_1, T_2 - \frac{d_2}{c_1}T_1) \) carries the same information as \((T_1, T_2)\). \( \square \)

Material omitted from Sect. 4

Proposition 3.

Proof Equation (10) follows immediately from the definition of \( AS \) together with the representation of the wage equation in (4) of Proposition 1. Since the trading event \( R < W \) can be written as \( AS + E[M | T] = E[R | T] - R \) and \( AS + E[M | T] \) and \( E[R | T] - R \) are independently distributed with \( E[R | T] - R \sim \mathcal{N}(0, \sigma_{R|T}^2) \), we have

\[ PR = \Pr \left[ AS + E[M | T] < \sigma_{R|T} \varepsilon | T \right] = 1 - F \left( \frac{AS + E[M | T]}{\sigma_{R|T}} \right). \]
Similarly,
\[
EC = E \left[ ((R - E[R | T]) - AS - E[M | T])^+ | T \right]
\]
\[
= \sigma_{R|T} E \left[ \left( \varepsilon - \frac{AS + E[M | T]}{\sigma_{R|T}} \right)^+ | T \right]
\]
and
\[
\Pi = E \left[ ((E[R | T] - R) + AS + E[M | T])^+ | T \right]
\]
\[
= \sigma_{R|T} E \left[ \left( \varepsilon + \frac{AS + E[M | T]}{\sigma_{R|T}} \right)^+ | T \right].
\]

The monotonicity statements follow immediately from these representations together with the fact that, regardless of the sign of \( \beta_{GR,T} \), \( AS + E[M | T] \) is increasing in \( E[M | T] \).

\[\Box\]

**Proposition 4.**

**Proof** Part 1 (a). It follows from (10) by the implicit function theorem and the fact that \( h' < 0 \) that \( AS \) is increasing point-wise in \( \beta_{GR,T} \). Similarly, \( AS \) is seen to be increasing in \( \sigma_{R|T} \) by the star-shaped property of the normal hazard function (see Lemma A1). For \( \Pi \), increasing \( \beta_{GR,T} \) affects \( \Pi \) only through its effect on \( AS \). \( AS \) has been seen to increase and therefore, by (13), so does \( \Pi \). Increasing \( \sigma_{R|T} \) has two effects. The indirect effect, through \( AS \), increases \( \Pi \) since \( \Pi \) is increasing in \( AS \) and \( AS \) is increasing \( \sigma_{R|T} \). That the direct effect is also positive follows easily from the representation \( \Pi = E[(\sigma_{R|T} \varepsilon + AS + E[M | T])^+ | T] \) and convexity of (.) +.

Part (b). The sign of the impact of \( \sigma_{E[M|T]} \) on \( AS \), is determined by the sign of \( \varepsilon \) in representation (14). Similarly for \( \Pi \). The convexity property of the normal hazard (see Lemma A1) is inherited by \( as \), hence it follows from Jensen’s inequality that \( E[AS] = E[as(E[M|T])] \) is increasing in \( \sigma_{E[M|T]} \). That expected profit increases with \( \sigma_{E[M|T]} \) follows from (13) and the fact that \( E[AS] \) and \( E[EC] \) are increasing in \( \sigma_{E[M|T]} \) (see Proof of 2(c) below).

Part 2(a). It follows from (10) that the sign of \( \frac{\partial}{\partial \sigma_{R|T}} \left( \frac{as + x}{\sigma_{R|T}} \right) \) is the same as the sign of \( \frac{x}{\sigma_{R|T}} \), that is of the opposite sign to \( x \) (which, recall here represents \( E[M | T] \)). Since \( PR \) is a decreasing function of \( \left( \frac{as + x}{\sigma_{R|T}} \right) \) the result follows. Similarly, since \( ec(x) = \sigma_{R|T} \Psi \left( \frac{x + as(x)}{\sigma_{R|T}} \right) \) with \( \Psi \) nonnegative decreasing, it follows that \( \frac{\partial ec(x)}{\partial \sigma_{R|T}} > 0 \) whenever \( x > 0 \).

Part (b). It follows from (12), since \( \Psi \) is a decreasing function and \( as \) is increasing in \( \beta_{GR,T} \), \( EC \) decreases in first-order stochastic dominance order as \( \beta_{GR,T} \) becomes larger. The same argument applies for \( PR \).

Part (c). Single-crossing follows since \( pr \) is a decreasing function of \( \sigma_{E[M|T]} \varepsilon \) and is therefore unchanged at \( \varepsilon = 0 \) but is increased by a decrease in \( \sigma_{E[M|T]} \) when \( \varepsilon > 0 \) and decreased when \( \varepsilon < 0 \). To establish that \( E[EC] \) is increasing in \( \sigma_{E[M|T]} \) is
more delicate. The difficulty lies in the fact that although \( \Psi \) and \( as \) are both convex, 
\[ x \mapsto \Psi \left( \frac{x + as(x)}{\sigma_{R|T}} \right) \] is not everywhere convex on account of \( \Psi \) being decreasing. Using the above representation, we can write \( E[EC] \) 
\[ E[EC] = \sigma_{R|T} E \left[ \Psi \left( \frac{\sigma_{E[M|T]}e + as \left( \frac{\sigma_{E[M|T]}e}{\sigma_{R|T}} \right)}{\sigma_{R|T}} \right) \right]. \] (19)

Differentiation gives 
\[ \frac{\partial E[EC]}{\partial \sigma_{E[M|T]}} = E \left[ \Psi' \left( \frac{1 - F}{1 - \beta_{GR,T}h'} \right) \right]. \]

Applying the implicit function theorem to the representation for \( as \) in (10) gives 
\[ as' = \frac{\beta_{GR,T}h'}{1 - \beta_{GR,T}h'}. \] (20)

Using this together with the fact that \( \Psi' = - \left( 1 - F \right) \) gives 
\[ \frac{\partial E[EC]}{\partial \sigma_{E[M|T]}} = -E \left[ \frac{1 - F}{1 - \beta_{GR,T}h'} \right]. \]

Setting \( k(x, \beta) = \frac{1 - F(x)}{1 - \beta h'(x)} \) and \( y(\epsilon) = \frac{\sigma_{E[M|T]}e + as(\sigma_{E[M|T]}e)}{\sigma_{R|T}} \), this becomes 
\[ \frac{\partial E[EC]}{\partial \sigma_{E[M|T]}} = -\int_{-\infty}^{\infty} k(y(\eta), \beta_{GR,T}) \eta f(\eta) d\eta. \] (21)

Using symmetry of the normal density around zero, 
\[ \frac{\partial E[EC]}{\partial \sigma_{E[M|T]}} = \int_{0}^{\infty} (k(y(-\eta), \beta_{GR,T}) - k(y(\eta), \beta_{GR,T})) \eta f(\eta) d\eta. \] (22)

The desired result now follows from the following facts: for each \( \beta < 1 \), (i) \( x \mapsto k(x, \beta) \) is quasi-concave; (ii) \( \frac{\partial}{\partial x} k(x, \beta) \big|_{x=0} = -\frac{\sqrt{\pi} (4\pi + \beta(2 + \pi))}{(\beta + 2\pi)^2} < 0 \) so \( x \mapsto k(x, \beta) \) is decreasing on the positive half-line; (iii) For \( x > 0, k(x, \beta) < k(-x, \beta); \) (iv) Hence, \( x \geq |y| \) implies \( k(x, \beta) < k(y, \beta). \) Since for \( \eta > 0, y(\eta) = \frac{\sigma_{E[M|T]}\eta + as(\sigma_{E[M|T]}\eta)}{\sigma_{R|T}} \) 
\[ \left| \frac{-\sigma_{E[M|T]}\eta + as(-\sigma_{E[M|T]}\beta)}{\sigma_{R|T}} \right| = |y(-\eta)|. \] It follows that for \( \beta < 1, \eta > 0, k(y(-\eta), \beta) > k(y(\eta), \beta) \) therefore immediately from the integration (21), \( \frac{\partial E[EC]}{\partial \sigma_{E[M|T]}} > 0 \) as required. \( \square \)
Proposition 5.

Proof Suppose \( \beta_{GR,T} \geq 0 \). Using the fact that \( h(x) \leq h(0) + (x)^+ \) in the representation (10) for \( as \):

\[
as(x) \leq \beta_{GR,T} \sigma_{R|T} h(0) + \beta_{GR,T}(as(x) + x)^+
\leq \beta_{GR,T} \sigma_{R|T} h(0) + \beta_{GR,T} as(x) + \beta_{GR,T}(x)^+,
\]

where the second inequality follows from the convexity of \( (x)^+ \). Collecting terms in \( as(x) \) and substituting the value of \( h(0) \) gives

\[
as(x) \leq \sigma_{R|T} \frac{\beta_{GR,T}}{1 - \beta_{GR,T}} \sqrt{\frac{2}{\pi}} + \frac{\beta_{GR,T}}{1 - \beta_{GR,T}} (x)^+.
\]

Similarly, noting \( h(x) \geq (x)^+ \) gives, from (10),

\[
as(x) \geq \beta_{GR,T} (as(x) + x)^+.
\]

Since \( as(x) \geq 0 \), if \( x \geq 0 \), then

\[
as(x) \geq \frac{\beta_{GR,T}}{1 - \beta_{GR,T}} x.
\]

Hence

\[
as(x) \geq \frac{\beta_{GR,T}}{1 - \beta_{GR,T}} (x)^+.
\]

These inequalities give the bounds in the Proposition. The \( \beta_{GR,T} \leq 0 \) case is dealt with by identical arguments. \( \square \)

Proposition 6.

Proof Parts (a) and (b) follow immediately from Proposition 4. To establish Parts (c) and (d), note that \( \text{Var}(W) = \text{Var}(E[G|T] - AS) = \sigma_{E[G|T]}^2 + \text{Var}(AS) - 2 \text{Cov}(AS, E[G|T]) \). By Stein’s Lemma

\[
\text{Var}(W) = \sigma_{E[G|T]}^2 + \text{Var}(AS) - E[as'(E[M|T])]2\text{Cov}(E[G|T], E[G|T]). \quad (22)
\]

Since \( \text{Var}(AS) \) and \( E[as'(E[M|T])] \) are independent of the parameters \( \sigma_{E[G|T]|E[M|T]} \) and \( \sigma_{E[G|T]} \), the results follow immediately from (22). \( \square \)
Material omitted from Sect. 5

Proposition 7.

Proof Observe that
\[ E[G | R < W] - E[G] = E[W | R < W] - E[G] \]
\[ = E[E[G|T] - E[G] | R < W] - E[AS | R < W] \]
\[ = \frac{\text{Cov}(E[G|T], PR)}{E[PR]} - \frac{E[AS.PR]}{E[PR]}, \]
where the first equality follows from the wage equation, the second by definition of adverse selection, and the third by elementary probability and definition of PR. Applying Stein’s Lemma to the first term and carrying out the multiplication AS × PR gives
\[ E[G | R < W] - E[G] = \frac{\beta_{GR.T} \sigma_{R|T} E[f]}{E[PR]} - \frac{\beta_{GR,T} \sigma_{R|T} E[f]}{E[PR]}, \]
where \( f \) and \( h' \) are respectively the \( \mathcal{N}(0, 1) \) density and derivative of hazard evaluated at \( \frac{AS + E[M|T]}{\sigma_{R|T}} \), and \( PR' \) is the derivative of \( pr \) evaluated at \( E[M|T] \). The conditions now follow upon noting that \( 0 < h' < 1 \). \( \square \)

Proposition 8.

Proof Applying the implicit function theorem to (10) and using the fact that \( h' < 1 \), one confirms that \( AS - \frac{\hat{f}_{GR,T}}{1 - \hat{f}_{GR,T}} E[M|T] \) is a decreasing function of \( E[M|T] \). Hence, since \( PR \) is also decreasing in \( E[M|T] \), it follows that \( \text{Cov} \left( AS - \frac{\hat{f}_{GR,T}}{1 - \hat{f}_{GR,T}} E[M|T], PR \right) \geq 0 \). Therefore, \( \text{Cov} (AS, PR) \geq \frac{\hat{f}_{GR,T}}{1 - \hat{f}_{GR,T}} \text{Cov} (E[M|T], PR) \). It follows that
\[ \text{Cov} (W, PR) \leq \text{Cov} \left( E[G|T] - \frac{\hat{f}_{GR,T}}{1 - \hat{f}_{GR,T}} E[M|T], PR \right). \]
Hence, since \( PR \) is decreasing in \( E[M|T] \) a sufficient condition for \( \text{Cov} (W, PR) \leq 0 \) is that \( \text{Cov} \left( E[G|T] - \frac{\hat{f}_{GR,T}}{1 - \hat{f}_{GR,T}} E[M|T], E[M|T] \right) \geq 0 \). Expanding this condition gives 2. A similar argument using the fact that \( AS \) is increasing in \( E[M|T] \) gives 1. \( \square \)
Material omitted from Sect. 6

Proposition 9.

Proof Part 1. Consider an information structure with symmetric information, i.e. with $T = (G, R)$. In any of the equilibria identified in Proposition 1(b), the allocation of the worker is efficient and the total surplus is $E[\max\{G, R\}]$. In the event that $G < R$, we may choose the equilibrium wage anywhere in the interval $(-\infty, R)$ (Proposition 1b). Making the selection, when $G < R$, of $W = R - \alpha$, for $\alpha \in [0, \infty)$ gives expected profit

$$E[(R - W)^+] = \Pr[R \geq G] > 0.$$ 

The strict inequality follows from our assumption $E[R] = E[G]$ which rules out $R < G$ with probability one. Hence, for appropriate choice of $\alpha$, expected profit $E[\Pi]$ may take any value in $[0, \infty)$ with $E[W] = E[\max\{G, R\}] - E[\Pi]$. The proof is completed by noting that by varying the quality of inside information, any value $E[\max\{G, R\}]$ in the interval $[0, E[\max\{V_G, V_R\}])$ may be attained.

Part 2. Lemma A2 below implies that for any any pair $(\beta_{GR,T}, \sigma_{E[M|T]})$ such that $\beta_{GR,T} < 1$ and $\sigma_M > \sigma_{E[M|T]} > \sigma_{E[M|R]} \geq 0$, we may select a sequence of information structures such that $(E[AS], E[EC])$ converges to $(E[AS], E[EC]) = \left(\frac{\beta_{GR,T}}{1 - \beta_{GR,T}}E[\left(\sigma_{E[M|T]}e\right)^{+}], E[\left(\sigma_{E[M|T]}e\right)^{+}]\right)$. Hence, given inside information, which determines $\sigma_{E[M|R]}$, any level of $E[EC]$ in the interval $\left(E[\left(\sigma_{E[M|R]}e\right)^{+}], E[\left(\sigma_M e\right)^{+}]\right)$ may be approached in the limit as $\sigma_{R|T} \rightarrow 0$. At the same time, for appropriate choice of $\beta_{GR|T} < 1$ any $E[AS]$ in the interval $(-E[\left(\sigma_{E[M|T]}e\right)^{+}], \infty)$ may be approached in the limit. This implies, using $E[W] = E[G] - E[AS]$ and $E[\Pi] = E[EC] + E[AS]$ that given inside information such that $\Sigma_{V_R V_G} = \Sigma_{RG} > 0$, by varying outside information appropriately one can approach in the limit any $(E[W], E[\Pi])$ in the set

$$\Omega_{GR} = \left\{(E[\Pi], E[W]) \in \mathbb{R}^2 : E[\Pi] \geq 0, E[V_R] \leq E[\Pi] + E[W] \leq E[\max\{G, R\}] \right\}.$$ 

Varying the quality of inside information across the interval $\Sigma_{V_R V_G} = \Sigma_{RG} > 0$, one obtains in the limit all possible $(E[W], E[\Pi])$ outcomes in $\Omega_{\Sigma_{V_R V_G}}$.

Part 3. Since by (12) of Proposition 3

$$EC = E[(W - R)^+] = \sigma_{R|T} \Psi\left(\frac{AS + E[M|T]}{\sigma_{R|T}}\right).$$

Moreover, by familiar properties of the $N(0, 1)$ distribution, $\Psi$ has the representation $\Psi(x) = f(x) - x(1 - F(x))$. Hence, writing $f$ for $f\left(\frac{AS + E[M|T]}{\sigma_{R|T}}\right)$ and similarly for
\[ EC = \sigma_{R|T} \left( f - \frac{f - AS + E[M|T]}{\sigma_{R|T}} (1 - F) \right). \]

Since \( EC \geq 0 \), multiplying by \( \frac{1}{1 - F} > 1 \) yields

\[ EC \leq \frac{EC}{1 - F} = \sigma_{R|T} \left( f - \frac{f - AS + E[M|T]}{\sigma_{R|T}} \right). \]

Using the representation (11) of Proposition 3

\[ EC \leq \left[ AS \frac{1 - \beta_{GR.T}}{\beta_{GR.T}} - E[M|T] \right]. \]

Implying upon taking expectations

\[ \mathbb{E}[EC] \leq \mathbb{E}[AS]\left[ \frac{1 - \beta_{GR.T}}{\beta_{GR.T}} \right]. \]

Adding \( \mathbb{E}[AS] \) to both sides and using the definition of profit (13), this gives

\[ \mathbb{E}[\Pi] \leq \mathbb{E}[AS]\left[ \frac{1}{\beta_{GR.T}} \right]. \]

Hence,

\[ \mathbb{E}[W] \leq \mathbb{E}[G] - f_{GR.T} \mathbb{E}[\Pi]. \]

\[ \qed \]

Lemma A2 (Existence of information structures with \( \sigma_{R|T} \to 0 \)). Suppose \( \sigma_{VR-VG} \geq \sigma_{M} > \sigma_{E[M|R]} \geq 0 \). For any pair \((s, b)\) such that \( \sigma_{M} > s > \sigma_{E[M|R]} \geq 0 \), and \( b < 1 \):

1. There exists an \( \epsilon > 0 \) such that for any \( 0 < r < \epsilon \), there is an information structure in \( \mathcal{P}^{+}_{\Sigma_{VR VG}} \) with \( (\beta_{GR.T}, \sigma_{E[M|T]}, \sigma_{R|T}) = (b, s, r) \).

2. As \( \sigma_{R|T} \) approaches zero,

   a. \( AS \) converges in distribution to \( \frac{\beta_{GR.T}}{1 - \beta_{GR.T}} (\sigma_{E[M|T]} \epsilon)^{+} \). Hence,

   \[ \lim_{\sigma_{R|T} \to 0} \mathbb{E}[AS] = \frac{\beta_{GR.T}}{1 - \beta_{GR.T}} \mathbb{E}\left[(\sigma_{E[M|T]} \epsilon)^{+}\right] = \frac{\beta_{GR.T}}{1 - \beta_{GR.T}} \sigma_{E[M|T]} f(0). \]

   b. \( \mathbb{E}[EC] \) converges to \( \lim_{\sigma_{R|T} \to 0} \mathbb{E}[EC] = \mathbb{E}\left[(\sigma_{E[M|T]} \epsilon)^{+}\right] = \sigma_{E[M|T]} f(0). \)

Proof Part 1. For \( \sigma_{R|T}^{2} > 0, \sigma_{M|T}^{2} = \sigma_{M}^{2} - \sigma_{E[M|T]}^{2} > 0, \Sigma_{GR|T} > 0 \) is equivalent to the condition

\[ \det \begin{pmatrix} \sigma_{R|T}^{2} & \sigma_{R|T}^{2}(1 - \beta_{GR.T}) \\ \sigma_{R|T}^{2}(1 - \beta_{GR.T}) & \sigma_{M}^{2} - \sigma_{E[M|T]}^{2} \end{pmatrix} > 0, \]
equivalently \( \left( \sigma^2_M - \sigma^2_{E[M|T]} \right) \geq \sigma^2_{R|T} (1 - \beta_{GR,T})^2 \). Given the assumed conditions on \( (\beta_{GR,T}, \sigma_{E[M|T]} \) this condition is satisfied in a neighbourhood of \( \sigma^2_{R|T} = 0 \). Similarly, \( \Sigma_{GR|T} \prec \Sigma_{GR} \) is equivalent to

\[
\det \begin{pmatrix}
\sigma^2_R - \sigma^2_{R|T} & \sigma_{RM} - \sigma^2_{R|T} (1 - \beta_{GR,T}) \\
\sigma_{RM} - \sigma^2_{R|T} (1 - \beta_{GR,T}) & \sigma^2_{E[M|T]}
\end{pmatrix} > 0,
\]

which as \( \sigma^2_{R|T} \to 0 \) becomes \( \det \begin{pmatrix}
\sigma^2_R & \sigma_{RM} \\
\sigma_{RM} & \sigma^2_{E[M|T]}
\end{pmatrix} > 0 \). That is, if \( \sigma^2_{E[M|T]} > \frac{\sigma^2_{RM}}{\sigma^2_R} = \sigma^2_{E[M|R]} \) there exists a neighborhood of \( \sigma_{R|T} = 0 \) in which \( \Sigma_{GR} \succ \Sigma_{GR|T} \succ 0 \).

Part 2(a). This is an immediate consequence of Proposition 5.

Part 2(b). Substitute the limiting value for adverse selection into the representation

\[
EC = \varepsilon \left( (\Delta R - AS - E[M|T])^+ | T \right),
\]

\[
\lim_{\sigma_{R|T} \to 0} \varepsilon \left( (\Delta R - AS - E[M|T])^+ | T \right) = \varepsilon \left( (1 - \beta_{GR,T}) \left( E[M|T])^+ - E[M|T] \right) \right).
\]

Note that \( \left( (1 - \beta_{GR,T}) \left( E[M|T])^+ - E[M|T] \right) \right) = \left( -\frac{\beta_{GR,T}}{1 - \beta_{GR,T}} (E[M|T])^+ - E[M|T] \right) = 0 \) whenever \( E[M|T] \geq 0 \) and is equal to \( (E[M|T])^+ \) whenever \( E[M|T] \leq 0 \), hence the desired result.

\( \square \)

**Material omitted from Sect. 7**

**Corollary 2.**

**Proof** It is immediate from (10) that \( \beta_{GR,T} < 0 \) is necessary and sufficient for advantageous selection. We can write

\[
\beta_{GR,T} = \frac{\text{Cov} \left( y - E[X|Q], y - E[X|Q] | T \right)}{\text{Var} \left( R|T \right)}.
\]

Since we assume that initial wealth \( y \) is non-random, the condition is simply that \( \text{Cov} \left( E[X|Q] + \frac{1}{2} A \sigma^2_{X|Q}, E[X|Q] | T \right) < 0 \). This inequality can be expressed as \( \sigma^2_{X|T} < -\frac{1}{2} \sigma^2_{X|Q} \sigma_{AX|T} \) which, upon re-arranging, gives the required result.

\( \square \)
References

Acemoglu, D., Pischke, J.-S.: Why do firms train? Theory and evidence. Q. J. Econ. **113**(1), 79–119 (1998)
Akerlof, G.: The market for ‘Lemons’: quality uncertainty and the market mechanism. Q. J. Econ. **84**(3), 488–500 (1970)
Arrow, K.: Essays in the Theory of Risk Bearing. Markham Pub. Co, Chicago (1971)
Bergemann, D., Morris, S.: Bayes correlated equilibrium and the comparison of information structures in games. Theor. Econ. **11**(2), 487–522 (2016)
Bergemann, D., Brooks, B., Morris, S.: The limits of price discrimination. Am. Econ. Rev. **105**(3), 921–957 (2015)
Birnbaum, Z.W.: An inequality for Mill’s ratio. Ann. Math. Stat. **13**(1), 245–246 (1942)
Birnbaum, Z.W.: Effect of linear truncation on a multinormal population. Ann. Math. Stat. **21**(2), 272–279 (1950)
Borjas, G.J.: Self-selection and the earnings of immigrants. Am. Econ. Rev. **77**(4), 531–553 (1987)
Chiappori, P.-A., Salanie, B.: Testing for asymmetric information in insurance markets. J. Polit. Econ. **108**(1), 56–78 (2000)
de Meza, D., Webb, D.: Advantageous selection in insurance markets. Rand J. Econ. **32**(2), 249–262 (2001)
Fang, H., Wu, Z.: Multidimensional private information, market structure and insurance markets. Rand J. Econ. **49**(3), 751–787 (2018)
Finkelstein, A., McGarry, K.: Multiple dimensions of private information: evidence from the long-term care insurance market. Am. Econ. Rev. **96**(4), 937–958 (2006)
Gibbons, R., Katz, L.: Layoffs and lemons. J. Labor Econ. **9**(4), 351–380 (1991)
Gordon, R.: Values of mill’s ratio of area to bounding ordinate of the normal probability integral for large values of the argument. Ann. Math. Stat. **12**(3), 364–366 (1941)
Greenwald, B.: Adverse selection in the labor market. Rev. Econ. Stud. **53**(3), 247–325 (1986)
Levin, J.: Information and the market for lemons. Rand J. Econ. **32**(4), 657–666 (2001)
Li, J.: Job mobility, wage dispersion, and technological change: an asymmetric information perspective. Eur. Econ. Rev. **60**, 105–126 (2013)
Marshall, A.W., Olkin, I.: Life Distributions. Springer, Berlin (2007)
Mas-Colell, A., Whinston, M.D., Green, J.R.: Microeconomic Theory. Oxford University Press, Oxford (1995)
Roy, A.: Some thoughts on the distribution of earnings. Oxford Econ. Pap. **3**(2), 135–146 (1951)
Sampford, M.R.: Some inequalities on mill’s ratio and related function. Ann. Math. Stat. **24**(1), 130–132 (1953)
Simpson, E.H.: The interpretation of interaction in contingency tables. J. Royal Stat. Soc. Ser. B **13**, 238–241 (1951)
Waldman, M.: Learning in Labor Markets. Edward Elgar, Broadheath (2017)
Yule, G.: Notes on the theory of association of attributes in statistics. Biometrika **2**, 121–134 (1903)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.