

doi: 10.1111/ijet.12079

Goethe's secret reserve price

Charles Z. Zheng*

The story in which Goethe sold a manuscript through a secret reserve price is revisited. Goethe's choice of secrecy can be explained as an attempt to mitigate the lemons problem through concealing his private information about the manuscript. The standard mechanisms that make the price public are unsafe for the seller as the buyer may adopt non-trusting posterior beliefs. A safe inscrutable mechanism is constructed here that generates more expected profits for some seller types than any such standard mechanism. A contrast between this mechanism and the one devised by Goethe explains why his mechanism was corrupted during its execution.

Key words auction theory, secret reserve price, informed principal, market for lemons, safe mechanisms, Goethe's auction, interdependent values

JEL classification D44, D82

Accepted 5 August 2015

1 Introduction

Goethe used a peculiar mechanism, involving a secret reserve price, to sell the manuscript of his epic poem "Hermann und Dorothea." In negotiating with the publisher Vieweg, Goethe wrote down a price in a sealed envelope and promised to sell the poem at the sealed price if Vieweg's bid turned out to be above it. This episode has been taken as Goethe's anticipation of the Vickrey auction, with his price in the sealed envelope corresponding to the reserve price in auction theory.¹ Given, however, that in a standard private-value auction model there is no gain from concealing the reserve price, why did Goethe conceal it?

Moldovanu and Tietzel (1998) have explained the concealment as Goethe's attempt to extract the publisher's private information for future use. Their explanation is based on a historical account of the book market back then, rather than on a formal model. Formalizing the explanation could be difficult, because the anticipation of future usage of the information could distort Vieweg's truthtelling incentive. Moreover, Goethe's mechanism failed during its execution. It turned out that a mediator, Counsel Böttiger, to whom Goethe entrusted the secret reserve price, revealed it to Vieweg before Vieweg submitted his bid. If values were really private according to this explanation, given the

^{*} Department of Economics, University of Western Ontario, London, Ontario, Canada. Email: charles.zheng@uwo.ca Financial support from the Social Science and Humanities Research Council of Canada is gratefully acknowledged.

¹ See, for example, Moldovanu and Tietzel (1998) and Lucking-Reiley (2000), as well as http://cheaptalk.org/2012/11/04/ the-goethe-auction/

standard risk-neutrality assumption in auction theory and the second-price-like payment scheme in Goethe's mechanism, Vieweg would have had nothing to gain from knowing Goethe's private information before submitting his bid.

Also with a private-value model, Yilankaya (1999) has demonstrated a case where some types of seller can profit from a mechanism which hides the seller's message that affects the buyer's payment amount. If the buyer naïvely bids according to his prior belief, the seller with sufficiently low use values of the good for sale would prefer a double auction, which conceals the seller's asking price until the end, to a mechanism that publishes the seller's asking price. While Yilankaya has shown that such gain from secrecy vanishes at equilibrium as a rational buyer would not maintain the prior belief upon seeing the seller's choice of the double auction, one could argue that Goethe's choice of secret reserve price might be an anticipation of Yilankaya's case subject to Goethe's understandable limitations on belief updating. Such an explanation, however, would have trouble reconciling with the difference that, conditional on a trade, the payment made by the buyer in Yilankaya's double auction depends on the buyer's bid, while not so in Goethe's mechanism. Had the payment rule in the double auction been changed to be independent of the buyer's bid, as in Goethe's auction, then in the private-value model Goethe's private information would not have mattered at all, nor would there have been any need for the mediator to reveal it, to the buyer.

Let us therefore provide an alternative explanation based on an idea that Goethe might have thought that what he privately knew could adversely affect a potential buyer's evaluation of the poem, leading to Goethe concealing his information in the hope of avoiding such an adverse effect. Specifically, consider the interdependence between Goethe's and the publisher's valuations of the manuscript, as well as the signaling effect of Goethe's sales offer. On the one hand, what Goethe knew about the quality of his poem would affect any publisher's evaluation of the publication rights of the poem. On the other hand, the publisher's knowledge about the book market would also affect Goethe's assessment of how much he could make from the poem. Thus, Goethe was facing Akerlof's (1970) lemons dilemma. His willingness to sell might be taken as a negative signal about the poem that could dampen a publisher's enthusiasm to buy it, and the publisher's willingness to pay a positive signal that could make Goethe reluctant to sell the poem right away.

Viewed from this perspective of mechanism design by a privately informed principal, Goethe's choice of hiding his reserve price can be understood as an attempt to mitigate the lemons problem. The standard alternative, a take-it-or-leave offer with a posted price, is *unsafe* in the sense that its performance depends on the posterior belief held by the buyer. The performance may be so bad that the seller may rather offer a mechanism that conceals his private information by keeping the price secret until the buyer has submitted a bid. Among such *inscrutable* mechanisms, those whose performances are independent of the buyer's posterior belief are called *safe*, a notion due to Myerson (1983). We shall demonstrate the gain from inscrutability by constructing a safe mechanism that would have generated more expected profits for Goethe for some types of private information. As in Goethe's mechanism, conditional on a trade, the buyer's payment is independent of the buyer's bid. We shall also show that Goethe's mechanism lacked a component necessary for it to be safe, which explains why his mechanism was eventually corrupted by the mediator.²

² Not only was Goethe back then hampered by his limited knowledge of game theory, but so also are we, as the literature has not reached the state of characterizing the equilibria of such an informed-principal game involving the lemons problem. For examples of the current literature, see Balkenborg and Makris (2015), who propose a novel solution concept in a finite-type, non-auction and common-value model where the buyer has no private information, and Mylovanov and Tröger (2012, 2014), who characterize an equilibrium refinement, strong neologism-proofness, in private-value models.

2 The gain from secrecy

Let us consider a simple setup of interdependent valuations. There are two players, both privately informed: a seller called player 0, our Goethe here, and a buyer called player 1, Vieweg in our story. Each player *i*'s signal is a random variable independently drawn from the uniform distribution on [0, 1]. If the realized signal is t_0 for the seller and t_1 for the buyer, the *ex post* value of the good (the publication rights of the poem) equals $t_0 + rt_1$ for the seller and $t_1 + rt_0$ for the buyer, where *r* is a parameter strictly between 0 and 1.

Suppose the two parties interact according to an informed-principal game, as first formulated by Myerson (1983). First, each player is privately informed of his realized signal, or *type*. Second, the seller publicly commits to a mechanism to sell an indivisible good. Third, the mechanism is operated to determine whether the good is sold and who pays whom how much. If the mechanism stipulates that the seller sends a message, then the seller does so accordingly. At this point, he can either choose a message to send or have it announced if he has anticipated his message beforehand, as Goethe did. Once the outcome mandated by the mechanism is carried out, the game ends.

A typical kind of selling mechanism in bilateral trades is the take-it-or-leave offer at a price announced by the seller at the outset. Such a mechanism with a posted price $p \in \mathbb{R}$, say, amounts to a communication game where the seller and buyer independently announce whether to accept the price; if both say yes then the good is traded at the price p, otherwise no trade occurs. With values interdependent, the seller's private signal affects the buyer's valuation of the good. Consequently, posted prices are *unsafe* in the sense that their outcomes depend on the buyer's posterior belief about the seller's signal inferred from the prices.

Following Maskin and Tirole (1990, 1992), let us *not* assume that the seller has the power of choosing an equilibrium in the game through persuading the buyer to adopt a posterior belief desirable to the seller. Then the seller, in choosing a mechanism, is faced with a situation similar to ambiguity aversion, with the role of multiple priors played here by the multiple posterior beliefs that the buyer may adopt. Thus, in the spirit of the maximin criterion for an ambiguity-averse decision-maker, let us focus on the *non-trusting posterior* belief that the seller's type is degenerate to the lowest possible point, zero.³ Suppose that, conditional on the seller's choice of any posted-price mechanism, the buyer adopts the non-trusting posterior. The seller's expected profits from posted-price mechanisms are then calculated.

Lemma 1 Given any posted price p, a type- t_1 buyer accepts p if and only if $t_1 \ge p$, and a type- t_0 seller's expected profit is equal to

$$\pi_0(t_0, p) := \begin{cases} \left(\left(1 - \frac{r}{2}\right) p - \frac{r}{2} - t_0 \right) (1 - p) & \text{if } t_0 \le (1 - r/2) p - r/2. \\ 0 & \text{if } t_0 \ge (1 - r/2) p - r/2. \end{cases}$$
(1)

PROOF: Based on the non-trusting posterior, a type- t_1 buyer's expected value of the good is equal to $t_1 + r \cdot 0 = t_1$, hence his best response is to accept the price p if and only if $t_1 \ge p$. Thus, the seller knows that, conditional on the event that the buyer accepts p, the buyer's type is uniformly distributed on [p, 1]. Hence any type- t_0 seller's expected value of the good is equal to $t_0 + r(p+1)/2$ conditional

³ Binmore *et al.* (2012) relate the history of how the maximin criterion came about. The opposite alternative to such a maximin perspective in the informed-principal literature is to consider equilibrium refinements such as Myerson's (1983) core and neutral mechanisms and Mylovanov and Tröger's (2012, 2014) strong neologism-proof mechanisms developed from Farrell (1993).

on the event that his decision is pivotal, that is, that the buyer accepts the price *p*. Consequently, his best response is to accept *p* if and only if $t_0 + r(p+1)/2 \le p$, that is, $t_0 \le (1 - r/2)p - r/2$, and Equation (1) follows.

Corollary 1 The posted price p^* that yields the highest expected profit for the type-0 seller among all posted prices is

$$p^* = \frac{1}{2 - r},$$
(2)

and

$$\pi_0(0, p^*) = \frac{(1-r)^2}{2(2-r)}.$$
(3)

PROOF: When $t_0 = 0$, Equation (1) becomes $\pi_0(0, p) = \max\left\{0, \left(1 - \frac{r}{2}\right)p - \frac{r}{2}\right\}(1-p)$. As a function of $p, \pi_0(0, p)$ is maximized at $p = 1/(2-r) =: p^*$. Then Equation (3) follows.

Instead of posting a price at the outset, the seller could keep it secret until the buyer's message has been sent. That was what Goethe did. More generally, the seller could use an *inscrutable mechanism* upon the announcement of which the seller provides no indication of his type; when operated, the mechanism solicits messages from both the buyer and the seller and determines the outcome based on their messages. To be credible from the buyer's perspective, an inscrutable mechanism needs to be *seller-incentive feasible* in the sense that truth-telling to the mechanism is the seller's best response based on his prior belief. (There is nothing for the seller at this point to update about the buyer, who has not acted.) According to Myerson (1983), such an inscrutable mechanism is called *safe* if it is not only seller-incentive feasible, but also buyer-incentive feasible whichever posterior belief that the buyer may hold. A safe mechanism provides a safeguard for the seller's profits because he can always count on its outcome, free of the tyranny of posterior beliefs. Furthermore, as demonstrated next, one can construct a safe mechanism that generates greater expected profits for a non-degenerate set of seller types.

Here is the intuition for such gain of secrecy. By the envelope-theorem characterization of incentive compatibility, the seller's expected profit at any equilibrium is a decreasing function of his type, and the rate of decrease equals the probability that trade occurs conditional on his type. In any posted-price mechanism, this probability is insensitive to the seller's type as long as the seller accepts the price. That is, once the seller has offered the good for sale at the posted price, he cannot send any further message to the mechanism. By contrast, in a safe inscrutable mechanism, the seller does not announce his type before the buyer's message is submitted, hence the probability decreases in the seller's type. Thus, as the seller's type increases, his expected profit decreases faster in a posted-price mechanism than in a safe mechanism, so eventually the former profit falls below the latter profit. That is a gain from secrecy, an insight that Goethe might have reached in opting for a secret reserve price.

3 A safe intrigue

Consider the following inscrutable mechanism, denoted \mathscr{S} . The seller and buyer simultaneously and independently submit reports of their types, say, t_0 from the seller and t_1 from the buyer. If $t_0 > 1 - \sqrt{r}$, then no trade occurs and neither party makes a payment. If $t_0 \le 1 - \sqrt{r}$, then the seller pays the buyer a signing bonus equal to

$$\frac{r(1-r)}{2(2-r)}t_0(2-t_0),\tag{4}$$

and if, in addition, $t_1 \ge \tau_1(t_0)$, where

$$\tau_1(t_0) := \frac{(1-r)t_0 + 1}{2-r},\tag{5}$$

then the good is sold to the buyer at a price equal to

$$\tau_1(t_0) + rt_0. \tag{6}$$

Lemma 2 Given any belief about the seller's type, \mathcal{S} is incentive feasible for the buyer.

PROOF: Pick any $t_0 \in [0, 1]$ and calculate the buyer's expected payoff conditional on the seller's type being t_0 . We are done if \mathscr{S} is incentive compatible and individually rational for the buyer conditional on t_0 . If $t_0 > 1 - \sqrt{r}$ then the buyer's payoff equals zero regardless of his actions. If $t_0 \le 1 - \sqrt{r}$, the buyer receives a signing bonus equal to expression (4), which is non-negative because 0 < r < 1 and $0 \le t_0 \le 1$ and is contingent only on t_0 ; in addition, the buyer receives an additional payoff equal to

$$t_1 + rt_0 - (\tau_1(t_0) + rt_0) = t_1 - \tau_1(t_0)$$

if and only if his report $\hat{t}_1 > \tau_1(t_0)$. Thus, it is optimal for the buyer to report truthfully, and participation guarantees him a non-negative payoff.

Lemma 3 *S* is incentive feasible for the seller (given the seller's prior belief).

PROOF: Let us calculate a type- t_0 seller's expected profit from reporting \hat{t}_0 . If $\hat{t}_0 > 1 - \sqrt{r}$ then he gets zero profit. If $\hat{t}_0 \le 1 - \sqrt{r}$ then the seller's expected profit equals

$$u_{0}(\hat{t}_{0}, t_{0}) = \left(1 - \tau_{1}(\hat{t}_{0})\right) \left(\tau_{1}(\hat{t}_{0}) + r\hat{t}_{0}\right) - \int_{\tau_{1}(\hat{t}_{0})}^{1} (t_{0} + rt_{1})dt_{1} - \frac{r(1 - r)}{2(2 - r)}\hat{t}_{0}(2 - \hat{t}_{0})$$

$$= \left(1 - \tau_{1}(\hat{t}_{0})\right) \left(\left(1 - \frac{r}{2}\right)\tau_{1}(\hat{t}_{0}) + r\hat{t}_{0} - t_{0} - \frac{r}{2}\right) - \frac{r(1 - r)}{2(2 - r)}\hat{t}_{0}(2 - \hat{t}_{0})$$

$$\stackrel{(5)}{=} \frac{1 - r}{2(2 - r)} \left((1 - \hat{t}_{0})\left((1 + r)\hat{t}_{0} - 2t_{0} + 1 - r\right) - r\hat{t}_{0}(2 - \hat{t}_{0})\right)$$

$$= \frac{1 - r}{2(2 - r)} \left((1 - t_{0})^{2} - r - (\hat{t}_{0} - t_{0})^{2}\right).$$
(7)

Thus, if $t_0 \le 1 - \sqrt{r}$ then $u_0(\cdot, t_0)$ is maximized at $\hat{t}_0 = t_0$ and the maximand is non-negative; if $t_0 > 1 - \sqrt{r}$ then $u_0(\cdot, t_0)$ is maximized at $\hat{t}_0 = 1 - \sqrt{r}$ and the maximum is negative, not as good

It follows from Lemmas 2 and 3 that the mechanism \mathscr{S} is safe. Next we show that \mathscr{S} beats any posted-price mechanism for some types of seller.

Proposition 1 For any posted-price mechanism, if the buyer adopts the non-trusting posterior belief, there exists a set of seller types, of strictly positive measure, who strictly prefer S to the posted-price mechanism.

PROOF: In \mathscr{S} , the probability of trade conditional on the seller's type t'_0 is equal to

$$q(t'_0) = \begin{cases} 0 & \text{if } t'_0 > 1 - \sqrt{r}, \\ 1 - \tau_1(t'_0) \stackrel{(5)}{=} \frac{1 - r}{2 - r} (1 - t'_0) & \text{if } t'_0 \le 1 - \sqrt{r}. \end{cases}$$
(8)

Denote by $U_0(t_0)$ a type- t_0 seller's expected profit from \mathscr{S} . By the envelope theorem (Milgrom and Segal 2002), due to the seller-incentive compatibility of \mathscr{S} ,

$$U_{0}(0) = U_{0}(1) + \int_{0}^{1} q(t_{0}')dt_{0}' \geq \int_{0}^{1} q(t_{0}')dt_{0}'$$

$$\stackrel{(8)}{=} \int_{0}^{1-\sqrt{r}} \frac{1-r}{2-r}(1-t_{0}')dt_{0}'$$

$$= \frac{1-r}{2-r}\left(1-\sqrt{r}-\frac{1}{2}\left(1-\sqrt{r}\right)^{2}\right) = \frac{(1-r)^{2}}{2(2-r)}$$

$$\stackrel{(3)}{=} \pi_{0}(0, p^{*}).$$
(9)

Recall from Lemma 1 that $\pi_0(t_0, p^*)$ is the type- t_0 seller's expected profit from the posted-price p^* that maximizes the type-0 seller's expected profit among all posted prices, given that the buyer adopts the non-trusting posterior, and Corollary 1 says that this optimal posted price for the type-0 seller is unique. Thus, there are only two possible cases:

- a. Either the given posted price p is not p^* . Then, by uniqueness of p^* , $\pi_0(0, p^*) > \pi_0(0, p)$. Then Inequality (9) implies that the type-0 seller strictly prefers the safe mechanism \mathscr{S} to the postedprice mechanism p. Furthermore, by continuity of U_0 and $\pi_0(\cdot, p)$, there is a non-degenerate interval $(0, \delta)$ such that $U_0(t_0) > \pi_0(t_0, p)$ for all $t_0 \in (0, \delta)$.
- b. Or the given posted price p is p^* . For any $t_0 \in [0, (1 r)/2]$, Equations (1) and (2) imply

$$\pi_0(t_0, p) = \pi_0(t_0, p^*) = \left(\left(1 - \frac{r}{2}\right) \frac{1}{2 - r} - \frac{r}{2} - t_0 \right) \left(1 - \frac{1}{2 - r}\right) = \frac{1 - r}{2(2 - r)}(1 - r - 2t_0).$$

Hence for any $t_0 \in (0, (1 - r)/2)$,

$$\frac{\partial}{\partial t_0}\pi_0(t_0, p) = -\frac{1-r}{2-r} < -\frac{1-r}{2-r}(1-t_0) = U_0'(t_0),$$

where the last equality follows from the envelope theorem and Equation (8). Thus, when t_0 increases from zero, $U_0(t_0)$ decreases more slowly than $\pi_0(t_0, p)$ does. This, coupled with the fact

 $U_0(0) \ge \pi_0(0, p)$ derived in inequality (9), implies that $U_0(t_0) > \pi_0(t_0, p)$ for all $t_0 \in (0, (1 - r)/2]$.

Therefore, in each case, there is a strictly positive measure of seller types who strictly prefer \mathscr{S} to the posted-price mechanism.

4 The urge for betrayal

While Goethe might have intuited the possible gain from secrecy illustrated above, it would be unfair to expect him back then to have been able to design a mechanism as elaborate as modern economic theory facilitates. Instead, his mechanism, denoted \mathscr{G} , amounts to soliciting a bid and an ask from the buyer and seller simultaneously so that trade occurs if and only if the buyer's bid is above the seller's ask, in which case the buyer pays for the good at a price equal to the seller's ask. While the mechanism is inscrutable and its payment scheme, as in our safe mechanism, makes the buyer's payment conditional on a trade independent of the buyer's bid, it is not safe.⁴

Proposition 2 *G* is not safe.

PROOF: Suppose that \mathscr{G} is safe, that is, there is a strategy profile $(s_0, s_1) : [0, 1]^2 \to \mathbb{R}^2$ that maps the type- t_0 seller to an ask $s_0(t_0)$ and the type- t_1 buyer to a bid $s_1(t_1)$ such that (i) for any $t_0 \in [0, 1]$, $s_0(t_0)$ is the type- t_0 seller's best reply to s_1 based on the prior distribution of the buyer's type, and (ii) for any $(t_0, t_1) \in [0, 1]^2$, $s_1(t_1)$ is the type- t_1 buyer's best reply to $s_0(t_0)$ provided that the seller's type is t_0 . Condition (ii) means that, for any $(t_0, t_1) \in [0, 1]^2$, $s_1(t_1)$ maximizes the type- t_1 buyer's *ex post* payoff $(t_1 + rt_0 - s_0(t_0)) \mathbf{1}_{x \ge s_0(t_0)}$ among all possible bids $x \in \mathbb{R}$. That implies $s_1(t_1) = t_1 + rt_0$ for all $(t_0, t_1) \in [0, 1]^2$, which is impossible because the buyer, in submitting his bid, does not know what t_0 is equal to.

Consequently, Goethe's mechanism is still at the mercy of the posterior belief held by the buyer, Vieweg, conditional on Goethe's choice of \mathscr{G} . In deciding on how much to bid in \mathscr{G} , Vieweg would be uncertain about Goethe's ask, which Vieweg wished to know in order to infer about Goethe's type, as the type affects Vieweg's valuation of the good. Whichever posterior belief he might have adopted, Vieweg's evaluation of the publication rights might fall below Goethe's asking price, resulting in no trade. In such an event, had he known what the ask was equal to and figured that the updated evaluation was above the ask, Vieweg would wish to have increased his bid.

This inefficiency problem of Goethe's mechanism was resolved by corruption. As mentioned earlier, Counsel Böttiger, to whom Goethe entrusted the secret reserve price, revealed it to Vieweg before Vieweg submitted his bid. Seeing Goethe's ask $s_0(t_0)$ thus provided an update about Goethe's type t_0 , and Vieweg could optimally bid up to his updated valuation of the publication rights. Consequently, given Goethe's ask $s_0(t_0)$, the outcome is efficient.

Counsel Böttiger's betrayal of Goethe's trust is rationalizable as his mediator role means that he could benefit, pecuniarily or not, from the gain of trade. Not every kind of inefficiency would result in such betrayal of confidence, however. For example, the safe mechanism \mathscr{S} is not efficient either, because trade occurs in \mathscr{S} only if $t_1 \ge \tau_1(t_0)$, that is, $(t_1 + rt_0) - (t_0 + rt_1) \ge 1 - t_1$. But had

⁴ An indirect mechanism such as \mathscr{G} is said to be *safe* if and only if it admits a Bayesian Nash equilibrium such that the outcome-equivalent direct revelation mechanism is safe.

Goethe used \mathscr{S} , there would be no point in Böttiger revealing Geothe's secret message to Vieweg, because by the definition of safeness Vieweg's action would not have been altered by the knowledge of Goethe's private information.

References

- Akerlof, G. A. (1970), "The market for 'lemons': Quality uncertainty and the market mechanism," *Quarterly Journal of Economics* 84, 488–500.
- Balkenborg, D., and M. Makris (2015), "An undominated mechanism for a class of informed principal problems with common values," *Journal of Economic Theory* **157**, 918–58.
- Binmore, K., L. Stewart, and A. Voorhoeve (2012), "How much ambiguity aversion? Finding indifferences between Ellsberg's risky and ambiguous bets," *Journal of Risk and Uncertainty* **45**, 215–38.
- Farrell, J. (1993), "Meaning and credibility in cheap-talk games," Games and Economic Behavior 5, 514-31.
- Lucking-Reiley, D. (2000), "Vickrey auctions in practice: From nineteenth century philately to twenty-first century e-commerce," *Journal of Economic Perspectives* 14, 183–92.
- Maskin, E., and J. Tirole (1990), "The principal-agent relationship with an informed principal: The case of private values," *Econometrica* 58, 379–409.
- Maskin, E., and J. Tirole (1992), "The principal-agent relationship with an informed principal, II: Common values," *Econometrica* **60**, 1–42.

Milgrom, P., and I. Segal (2002), "Envelope theorems for arbitrary choice sets," Econometrica 70, 583-601.

- Moldovanu, B., and M. Tietzel (1998), "Goethe's second-price auction," Journal of Political Economy 106, 854-9.
- Myerson, R. B. (1983), "Mechanism design by an informed principal," Econometrica 51, 1767-97.
- Mylovanov, T., and T. Tröger (2012), "Informed-principal problems in environments with generalized private values," *Theoretical Economics* 7, 465–88.
- Mylovanov, T., and T. Tröger (2014), "Mechanism design by an informed principal: Private values with transferable utility," *Review of Economic Studies* **81**, 1668–707.
- Yilankaya, O. (1999), "A note on the seller's optimal mechanism in bilateral trade with two-sided incomplete information," *Journal of Economic Theory* 87, 267–71.