THE ROLE OF THE AGENT'S OUTSIDE OPTIONS IN PRINCIPAL-AGENT RELATIONSHIPS*

IMRAN RASUL†
UNIVERSITY COLLEGE LONDON

SILVIA SONDEREGGER‡
UNIVERSITY OF BRISTOL AND CMPO

OCTOBER 2009

Abstract

We consider a principal-agent model of adverse selection where, in order to trade with the principal, the agent must undertake a relationship-specific investment which affects his outside option to trade, i.e. the payoff that he can obtain by trading with an alternative principal. This creates a distinction between the agent's *ex ante* (before investment) and *ex post* (after investment) outside options to trade. We investigate the consequences of this distinction, and show that whenever an agent's *ex ante* and *ex post* outside options differ, this may equip the principal with an additional tool for screening among different agent types, by randomizing over the probability with which trade occurs once the agent has undertaken the investment. In turn, this may enhance the efficiency of the optimal second-best contract.

JEL Classification: D21, L14.

Keywords: adverse selection, randomization, type-dependent outside options.

*We thank the associate editor and two anonymous referees for suggestions that have helped improve the paper. We have also benefitted from the comments of Fabrizio Adriani, Roman Inderst, Ian Jewitt, Bruno Jullien, Thomas Mariotti, In-Uck Park, Francesco Squintani, Max Steuer, Thomas von Ungern-Sternberg and seminar participants at LSE and the University of Lausanne. All errors are our own.

†Department of Economics, University College London, Drayton House, 30 Gordon Street, London WC1E 6BT, United Kingdom. Tel: +44-207-679-5853; E-mail: i.rasul@ucl.ac.uk.

‡Department of Economics, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, United Kingdom. Tel: +44-117-928-9039; E-mail: s.sonderegger@bristol.ac.uk.

1 Introduction

In many forms of bilateral exchange, one party has to undertake relationship-specific investments before trade can occur with their partner. An important consequence of such specific investments is that they typically change the investing party's outside option to trade, namely the payoff that he would obtain by trading with an alternative partner. For example, a firm that tailors its machinery in order to produce a specific widget required by a certain buyer, will change its production possibilities when trading with alternative buyers whose requirements need not be the same.[1]

A key distinction therefore exists between the firm's *ex ante* outside option, before the relationship-specific investment is undertaken, and their *ex post* outside option, after the investment has occurred. This paper investigates the consequences of this distinction in principal-agent models of adverse selection, where the agent's type is his private information, and both parties are risk neutral. We show that whenever an agent's *ex ante* and *ex post* outside options differ, this may equip the principal with an additional tool for screening among different agent types, by randomizing over the probability with which trade occurs once the agent has undertaken the specific investment. In turn, this may enhance the efficiency of the optimal second-best contracts.

This paper contributes to the literature on mechanism design when agents have type-dependent outside options (Lewis and Sappington 1989, Maggi and Rodriguez-Clare 1995, Jullien 2000). The literature on adverse selection identifies several cases in which the optimal mechanism can involve randomization, such as when agents have different levels of risk aversion (Stiglitz 1982, Arnott and Stiglitz 1988, Brito *et al* 1995), when the agent's type-space is multi-dimensional (Baron and Myerson 1982, Rochet 1984 and Thanassoulis 2004), or when randomization might allow non-monotonic allocation schedules to become incentive compatible (Strausz 2006). A further rationale for randomization is presented by Calzolari and Pavan (2006), who show that, in principal-agent problems with sequential contracting, randomization may be optimal, since it allows one principal to hide information from another principal. We add to the literature by considering situations where relationship-specific investments affect the agent's future prospects, so that his type-dependent *ex ante* and *ex post* outside options differ. This provides a novel rationale of why randomization may be optimal in principal-agent settings.

The remainder of the paper is organized as follows. In Section 2 we develop the principal-agent model. Section 3 solves for the optimal second best contracts. Section 4 discusses the efficiency consequences of having both types of outside option and also addresses possible extensions. All proofs and a numerical example are in the Appendix.

2 Model

Preliminaries We consider a principal-agent model with a principal P and an agent A, who contract over the production of output, q. Production is assumed to be observable and verifiable. The agent's marginal cost of production, θ, which defines his type, is not observed by the principal, and we assume

[1] This phenomenon is not confined to bilateral exchange between firms. Consider a traveller who wants to travel from A to B at 8pm on a given day. The traveller can choose whether to travel by train or bus. The specific investment undertaken by the traveller in order to access a certain type of travel takes the form of him being physically present at a particular location – the bus or train station – at a particular time. While from an *ex ante* perspective the traveller's outside option to catching the 8pm bus would be to take the 8pm train, once he has made the specific investment of arriving at the bus station prior to 8pm, his *ex post* outside option to catching the 8pm bus will be quite different. While he may for example catch the 9pm train, the 8pm train has been ruled infeasible by his earlier specific investment.

$\theta \in \{\theta_H, \theta_L\}$, where $\theta_H > \theta_L > 0$, and $prob(\theta = \theta_H) = \lambda$. In order to trade with the principal, the agent must undertake a relationship-specific investment, with cost normalized to zero. The agent's decision to undertake the investment is observable and verifiable. A contract between the principal and the agent is denoted $\{\phi, \pi, q, T\}$, where $\phi \in \{0, 1\}$ specifies whether the agent must undertake the investment[2], $\pi \in [0, 1]$ denotes the probability with which trade occurs between the parties, $q \in [0, \overline{q}]$ denotes the output that the agent must produce in case of trade, and $T \in \mathbb{R}^+$ indicates the payment from the principal to the agent (independent of whether trade actually occurs or not). We assume trade can only occur if the agent has made the relationship-specific investment so that if $\phi = 0$, $\pi = 0$.[3]

The principal's problem consists of designing the optimal menu of contracts from which the agent makes his preferred choice. The revelation principle states this search can be confined to the set of direct revelation mechanisms, whereby the agent is requested to report his type and is offered a contract that is contingent upon this report. The timing of actions is then as follows.

t=0 P offers A a menu of contracts $M = \{M_H, M_L\}$, where $M_i = \{\phi_i, \pi_i, q_i, T_i\}$ is the contract offered to the agent when his reported type is θ_i, $i = H, L$.

t=0.5 If A accepts M_i and M_i specifies $\phi_i = 1$, A undertakes the relationship-specific investment.

t=1 Conditional on $\phi_i = 1$, trade occurs with probability π_i, in which case A produces q_i. With probability $1 - \pi_i$ trade between A and P does not occur. If $\phi_i = 0$, trade between A and P does not occur with certainty.

t=1.5 Provided that he has respected the terms of the contract, A receives T_i.

Without loss of generality we restrict attention to contracts that always induce truthtelling and participation by the agent.

Agent's *Ex ante* and *Ex post* Outside Options If the agent does not accept the principal's contract, or if his contract prescribes $\phi_i = 0$, then the agent does not undertake any relationship-specific investment, and obtains a payoff $B_i \geq 0$ from alternative trade, where $i = H, L$. This defines the agent's *ex ante* outside option. Importantly, we allow for the possibility that *ex ante* outside options differ across types, so that $B_H \neq B_L$. If the agent undertakes the relationship-specific investment, but trade between the parties does not occur, then the agent obtains a payoff $C_i < B_i$ from alternative trade. C_i captures the agent's *ex post* outside option, namely the value of his trading prospects with alternative principals, after having undertaken the relationship-specific investment with the previous principal. *Ex post* outside options may also be type-dependent, so that $C_H \neq C_L$. The expression $B_i - C_i > 0$ reflects the loss in terms of the agent's alternative trading prospects from undertaking the relationship-specific investment, which tailors his production to the principal's needs. We refer to this as the opportunity cost of randomization, since this cost is only incurred when $\pi_i < 1$.

Payoffs Both parties are assumed to be risk neutral with respect to monetary transfers and production. If a type θ_i agent accepts a contract $\{\phi_i, \pi_i, q_i, T_i\}$, his net expected utility is,

$$u(\theta_i) = T_i + \phi_i \{-\theta_i \pi_i q_i + (1 - \pi_i) C_i - B_i\}. \tag{1}$$

[2]Allowing the contract to specify ϕ enables us to restrict attention to contracts that are always accepted by the agent. We thank an anonymous referee for providing this suggestion.

[3]By restricting attention to $\phi \in \{0, 1\}$ we rule out the possibility of the principal randomizing over ϕ. This is done to shorten the exposition of our results. The possibility of randomization over ϕ is discussed in section 4.

The principal's expected payoff is $U_P = \phi_i \pi_i v q_i - T_i$, where $v > \theta_H$. Let u_i denote the utility obtained by a type θ_i agent when he truthfully declares his type. From (1), holding constant all other dimensions of the contract offered to type θ_i, there is a one-to-one relation between T_i and u_i. In what follows we will therefore characterize a contract as $M_i = \{\phi_i, \pi_i, q_i, u_i\}$. Finally, we denote $\theta_H - \theta_L$ as $\Delta\theta$, $C_H - C_L$ as ΔC, $B_H - B_L$ as ΔB and $u_H - u_L$ as Δu.

3 Results

The participation constraint for a type θ_i agent is $u_i = T_i + \phi_i [-\theta_i \pi_i q_i + (1-\pi_i)C_i - B_i] \geq 0$. The incentive compatibility constraints which ensure agents find it optimal to declare their true type are,

$$\begin{aligned} IC_H &: u_H \geq u_L + \phi_L [-\pi_L q_L \Delta\theta + (1-\pi_L)\Delta C - \Delta B] . \\ IC_L &: u_L \geq u_H + \phi_H [\pi_H q_H \Delta\theta - (1-\pi_H)\Delta C + \Delta B] . \end{aligned}$$

Suppose full information contracts are offered so that $\phi_i = \pi_i = 1$, $q_i = \overline{q}$, and $u_i = 0$ for $i = H, L$. Constraint IC_H becomes, $0 \geq -\overline{q}\Delta\theta - \Delta B$, and IC_L becomes, $0 \geq \overline{q}\Delta\theta + \Delta B$. We focus on the more intuitive case in which $\overline{q}\Delta\theta + \Delta B > 0$ so that θ_L types have incentives to overstate their costs and mimic θ_H types. This is embodied in assumption A1 below.[4] To ensure that under full information the optimal contract prescribes $\phi_i = \pi_i = 1$, $q_i = \overline{q}$ for both types, assumption A2 below is required.

A1: $\overline{q}\Delta\theta + \Delta B > 0$

A2: $\overline{q}(v - \theta_i) \geq B_i$, $i = H, L$

Our first result provides a partial characterization of type θ_H's optimal contract whenever θ_H agents are required to undertake the relationship-specific investment.

Lemma 1: *It is never optimal for the principal to offer $\phi_H = 1$ in conjunction with π_H and q_H satisfying,*

$$\pi_H q_H \Delta\theta - (1-\pi_H)\Delta C + \Delta B < 0. \tag{2}$$

Under A1 the full information contracts would violate IC_L. By offering type θ_H agents a contract such that $\pi_H q_H \Delta\theta - (1-\pi_H)\Delta C + \Delta B = 0$, the principal ensures both that IC_L is satisfied *and* that no rents are offered to θ_L agents. Offering θ_H agents a contract such that (2) holds would only increase the distortions of π_H and/or q_H from their full information values (1 and $\overline{q}$ respectively) without generating any gain for the principal. This is essentially the rationale for Lemma 1.

An implication of Lemma 1 is that the participation constraint of type θ_L will not bind at the optimum. This is because, given type θ_H's participation, IC_L implies $u_L \geq u_H \geq 0$. In what follows, we therefore allow IC_L to hold with equality, let $u_H = 0$, and ignore constraint IC_H. We then later verify that the solution of the relaxed problem indeed satisfies IC_H. The principal's problem then is,

$$\max_{\substack{q_i \in [0,\overline{q}],\ \pi_i \in [0,1], \\ \phi_i \in \{0,1\},\ i=H,L}} U_P = \begin{array}{l} \lambda\phi_H [\pi_H q_H (v - \theta_H) + (1-\pi_H)C_H - B_H] + \\ (1-\lambda)\phi_L [\pi_L q_L (v-\theta_L) + (1-\pi_L)C_L - B_L] - \\ (1-\lambda)\phi_H [\pi_H q_H \Delta\theta - (1-\pi_H)\Delta C + \Delta B] \end{array} \tag{$\mathbf{P}$}$$

[4] For completeness, in the Appendix, we state the main results for the case in which the parameter values are such that high types have incentives to understate their type and mimic low cost types. These two cases arise because of the existence of the type-dependent *ex ante* outside options, B_i, as has been analyzed in detail by Maggi and Rodriguez-Clare (1995). Note that in the knife-edge case where $\overline{q}\Delta\theta + \Delta B = 0$ the principal can offer the full information contract to both types without inducing either to mimic the other, so this is clearly her favored course of action.

$$\text{subject to} \quad \phi_H \left[\pi_H q_H \Delta\theta - (1-\pi_H)\Delta C + \Delta B\right] \geq 0. \qquad \textbf{(C1)}$$

where (**C1**) derives from Lemma 1. We first solve (**P**) ignoring (**C1**). If the solution satisfies (**C1**) with strict inequality, it is the solution to the overall problem. Otherwise (**C1**) binds.

The principal faces a standard trade-off between efficiency and informational rents. If she offers θ_H types the efficient (full-information) contract where $\phi_H = \pi_H = 1$, $q_H = \overline{q}$, then she must also offer positive rents to θ_L types to prevent mimicking. In this case (**C1**) is slack. If the principal wishes to eliminate θ_L's rents, then she must distort type θ_H's contract away from the efficient contract.[5] In this case (**C1**) binds so, conditional on $\phi_H = 1$, we have,

$$q_H = \frac{(1-\pi_H)\Delta C - \Delta B}{\pi_H \Delta\theta}. \qquad (3)$$

When $\Delta C - \Delta B > 0$ – i.e., the opportunity cost of randomization is higher for θ_L than for θ_H – then (3) implies that $\partial q_H / \partial \pi_H < 0$. By lowering π_H the principal can increase q_H whilst keeping θ_L's rents at zero. A trade-off then emerges. A lower π_H decreases the probability of trade, but it also increases q_H, and hence the value of trade. When the latter effect is stronger than the former, the optimal contract (conditional on **C1** binding) prescribes $q_H = \overline{q}$ and $\pi_H = (\Delta C - \Delta B) / (\overline{q}\Delta\theta + \Delta C) \in (0,1)$, i.e. it prescribes randomization.[6] Proposition 1 fully describes the optimal second best contracts.[7]

Proposition 1: *For type θ_L, the optimal contract always prescribes $\phi_L = \pi_L = 1$, $q_L = \overline{q}$. If*

$$\lambda > \max\left\{\frac{\Delta C + \overline{q}\Delta\theta}{\overline{q}\,(v-\theta_L) - C_L}, \frac{\Delta\theta}{v-\theta_L}, \frac{\Delta B + \overline{q}\Delta\theta}{\overline{q}\,(v-\theta_L) - B_L}\right\}, \qquad (4)$$

*then (**C1**) is slack, and the optimal contract for type θ_H has $\phi_H = \pi_H = 1$, $q_H = \overline{q}$. If (4) does not hold, then (**C1**) binds, and the optimal contract for type θ_H is*

(i) if $C_H > \frac{C_L(v-\theta_H)}{v-\theta_L}$ and $\Delta C - \Delta B > \frac{(B_H - C_H)(\overline{q}\Delta\theta + \Delta C)}{\overline{q}(v-\theta_H) - C_H} > 0$: $\phi_H = 1$, $\pi_H = \frac{\Delta C - \Delta B}{\overline{q}\Delta\theta + \Delta C}$ and $q_H = \overline{q}$.

(ii) if $C_H < \frac{C_L(v-\theta_H)}{v-\theta_L}$ and $\Delta B < -\frac{B_H \Delta\theta}{v-\theta_H} < 0$: $\phi_H = \pi_H = 1$ and $q_H = -\frac{\Delta B}{\Delta\theta}$.

(iii) in all the other cases: $\phi_H = 0$.

If *prob*$(\theta = \theta_H) = \lambda$ is sufficiently high, then the principal finds it optimal to offer θ_H types the efficient contract, so as to maximize her profit when trading with θ_H types, even if this implies that positive rents are relinquished to agents of type θ_L. Conversely, if λ is sufficiently low, then the principal prefers to allow (**C1**) to bind and so eliminate any rents to θ_L types.

Proposition 1 shows that in order for the optimal contract for θ_H to prescribe randomization, $\Delta C - \Delta B$ should be positive, and sufficiently large. Intuitively, $\Delta C - \Delta B > 0$ implies that the opportunity cost of randomization is higher for θ_L types than for θ_H. Hence, by offering θ_H types a contract involving randomization, the principal can lower the incentives of θ_L types to overstate their costs and mimic θ_H types. By contrast, if $\Delta C - \Delta B < 0$, then θ_H types stand to lose more from randomization than θ_L types, and so randomization would not help deter θ_L from mimicking θ_H. Similarly, if $\Delta C = \Delta B = 0$ – as is the case if both *ex ante* and *ex post* outside options are type-invariant so $B_H = B_L$ and $C_H = C_L$ – then the opportunity cost of randomization is the same for both types, and again randomization is not an

[5] Given the linearity of her payoff, the principal would never select contracts between these extremes.

[6] That $(\Delta C - \Delta B) / (\overline{q}\Delta\theta + \Delta C) \in (0,1)$ follows from $\Delta C - \Delta B > 0$ and assumption A1.

[7] We adopt the convention that if P is indifferent between setting $\phi_i = 1$ or $\phi_i = 0$ for $i = H, L$, then she selects $\phi_i = 0$. Similarly, if she is indifferent between all $\pi_i \in [0,1]$ (resp., all $q_i \in [0, \overline{q}]$) then P selects $\pi_i = 0$ (resp., $q_i = 0$).

effective screening tool. This clarifies why type-dependent outside options are essential for randomization to be optimal.

Note that, in order for randomization to be optimal, $\Delta C - \Delta B$ should not only be positive, but also sufficiently large. This ensures that a small amount of randomization in the contract offered to θ_H is sufficient to deter θ_L from mimicking, and guarantees that the principal can obtain a positive expected profit when trading with type θ_H.

What are the implications of $\Delta C - \Delta B > 0$? From (3), we know that when (**C1**) binds and $\Delta C - \Delta B > 0$ then a trade-off emerges between π_H and q_H. A lower π_H decreases the probability of trade, but it also increases q_H, and hence the value of trade. For randomization to be optimal, the principal must then be willing to lower the probability of trade with θ_H in order to raise q_H at the margin. Whether this occurs or not, depends on the precise comparison between the costs (i.e., trade with θ_H occurs less often) and the benefits (i.e., q_H is higher) of randomization.

To see how the former may outweigh the latter, consider the simple case where $C_L < C_H < 0$, $B_H = B_L = 0$. Since *ex ante* outside options are independent of type, if $\pi_H = 1$, this case corresponds to the canonical model. As can be seen from (**C1**), leaving no rents to type θ_L then requires $q_H = 0$. The principal's payoff when dealing with type θ_H is then equal to zero. By contrast, setting $\pi_H < 1$ allows the principal to set $q_H = (1 - \pi_H) \Delta C / \pi_H \Delta\theta > 0$. Here, the cost of imposing randomization is null, since when $\pi_H = 1$ trading with type θ_H generates no profits (this follows from $q_H = 0$). By contrast, if C_H is not too negative, i.e. $C_H > C_L(v - \theta_H) / (v - \theta_L)$, then the benefit of randomization is strictly positive, since it allows the principal to obtain a strictly positive expected payoff when dealing with θ_H.[8] Conditional on the principal wishing to leave no rents to type θ_L (which, as highlighted by proposition 1, happens whenever λ is sufficiently low), randomization is then clearly optimal. Given the linearity of her payoff, if the principal finds it optimal to sacrifice π_H in order to raise q_H at the margin, then she goes all the way, and sets q_H as high as possible in the optimal contract, i.e. $q_H = \overline{q}$. From (3), π_H is then equal to $(\Delta C - \Delta B) / (\overline{q}\Delta\theta + \Delta C)$.

A Numerical Example In the Appendix, we discuss a numerical example where $\theta_H = 0.75$, $\theta_L = 0.25$, $\overline{q} = v = 2$, and the agent's *ex ante* and *ex post* outside options are $B_H = 1.85$, $B_L = 2.35$, $C_H = 1.75$, and $C_L = 1.95$. In that case, it is straightforward to show that, for $\lambda < 0.52$, the optimal contract offered to type θ_H prescribes $\phi_H = 1, \pi_H = 0.375$ and $q_H = \overline{q} = 2$.

4 Discussion

Efficiency Proposition 1 highlights the impact of having *two* (i.e., *ex ante* and *ex post*) type-dependent outside options on the optimal second best contracts. Suppose that, on the contrary, $C_i = B_i$ for both $i = H, L$, so $\Delta C = \Delta B$. From (3), the only way for (**C1**) to then bind is to set $q_H = -\Delta B / \Delta\theta$. If (4) does not hold and $\Delta B \geq -B_H \Delta\theta / (v - \theta_H)$, then the optimal contract prescribes $\phi_H = 0$, i.e. no trade between the principal and agents of type θ_H, since with $q_H = -\Delta B / \Delta\theta$ the principal would never

[8] The principal's expected payoff from dealing with θ_H when $q_H = (1 - \pi_H) \Delta C / \pi_H \Delta\theta$ is equal to $(1 - \pi_H) [C_H(v - \theta_L) - C_L(v - \theta_H)] / \Delta\theta - B_H$. In this simple example, $B_H = 0$. The condition $C_H > C_L(v - \theta_H) / (v - \theta_L)$ is therefore both necessary and sufficient to ensure that randomization is optimal whenever (**C1**) binds. More generally, this requirement is only necessary (as highlighte in part (i) of proposition 1, another condition is also required). Intuitively, if C_H is very low, then the transfer necessary to induce type θ_H to accept a contract involving randomization would be large, and randomization would therefore not be optimal.

obtain a non-negative profit when dealing with type θ_H. In contrast when $B_i \neq C_i$, trade between the principal and agents of type θ_H may occur with positive probability even if $\Delta B \geq -B_H \Delta\theta / (v - \theta_H)$.[9]

Hence, in a complete contracting environment, the need for agents to undertake relationship-specific investments *ex ante* that decrease the agent's outside option, can result in greater *ex post* efficiency, that is, at the production stage. This is because such investments enable the principal to utilize randomization as a tool to screen between agent types. To our knowledge, the earlier literature has not noted this potentially useful role for *ex ante* relationship-specific investments to improve on *ex post* efficiency. The literature has emphasized rather, that in the presence of contractual incompleteness, investment specificity results in *ex ante* inefficiencies, i.e. inefficiencies at the investment stage (Grout 1984, Grossman and Hart 1986, Hart and Moore 1990).

Relaxing the Linearity Assumption The restriction to linear payoff functions allows us to abstract from risk-aversion considerations, and to differentiate our results from the existing literature on randomization in mechanism design (Stiglitz 1982, Arnott and Stiglitz 1988, Brito *et al* 1995). However, our results extend also to non-linear settings. To see this, suppose agents face convex production costs, so the net utility of an type θ_i agent when accepting a contract $\{\phi, \pi, q, T\}$ is,

$$T + \phi \left[-\pi\theta_i g(q) + (1-\pi)C_i - B_i \right]. \tag{5}$$

where $g'(q) > 0$ and $g''(q) > 0$ for all $q > 0$. Suppose that, if offered the full-information contract, a type θ_L agent would overstate his cost and mimic type θ_H, as was the case throughout Section 3. Condition (**C1**) then is,

$$\phi_H \left[\pi_H g\left(q_H\right) \Delta\theta - (1-\pi_H)\Delta C + \Delta B \right] \geq 0. \tag{\textbf{C1}$'$}$$

Following the same argument as in Proposition 1, for λ sufficiently low, the optimal contract for type θ_H agents is such that **C1′** binds. Then, conditional on $\phi_H = 1$, we have,

$$q_H = g^{-1}\left(\frac{(1-\pi_H)\Delta C - \Delta B}{\pi_H \Delta\theta} \right). \tag{6}$$

As in the linear case, when $\Delta C - \Delta B > 0$, (6) implies that $\partial q_H / \partial \pi_H < 0$, so that, by lowering π_H, the principal can increase q_H. Clearly, this is a necessary requirement for randomization to be offered, or else the principal would always optimally select $\pi_H = 1$. Note that, in contrast with the linear case above, in this non-linear case the trade-off between π_H and q_H may actually make U_P concave in π_H – thus warranting randomization even in the absence of restrictions on feasible output quantities. To see this, let $g(q) = 0.5q^2$, and suppose that parameter values continue to follow the numerical example given above, but the restriction that q may not exceed $\overline{q}$ is relaxed.

Expression (6) then becomes $q_H = \sqrt{0.8 + \frac{1.2}{\pi_H}}$. Conditional on $\phi_H = 1$, the principal's expected payoff when dealing with a type θ_H agent is $U_P = \pi_H \left[2\sqrt{0.8 + \frac{1.2}{\pi_H}} - 0.75\left(0.4 + \frac{0.6}{\pi_H}\right) - 1.75 \right] - 0.1$, which is concave in π_H. The optimal contract for θ_H is $\phi_H = 1$, $\pi_H = 0.78$, and $q_H = 1.53$, and when dealing with type θ_H, the principal's expected profit is 0.23. Hence in this numerical example, for λ sufficiently low the optimal contract for θ_H may again prescribe randomization, although in contrast with the linear case, the optimal q_H is below its first-best value.

[9] This is the case for instance in the numerical example introduced above, where $\Delta B = -0.5 > -B_H \Delta\theta / (v - \theta_H) = -0.74$.

Allowing for Randomization Over ϕ What would happen if instead of assuming $\phi \in \{0,1\}$, ϕ were allowed to take any value in $[0,1]$? In that case, randomization over ϕ_H would be possible. However, we argue that it would not be optimal.[10] To see why, note that, similar to what happens for randomization over π_H, randomization over ϕ_H may only be optimal when (**C1**) binds – ϕ_H enters the principal's payoff linearly, so if (**C1**) is slack then ϕ_H optimally takes a corner value. Consider now a contract $M_H = \{\phi_H, \pi_H, q_H, u_H\}$, where $\phi_H \in (0,1)$ and $u_H = 0$ (as discussed above, this latter condition is always satisfied at an optimum). Under this contract, the opportunity cost of randomization over ϕ_H for an agent of type θ_i is equal to $-\theta_i \pi_H q_H + (1-\pi_H)C_i - B_i$, namely the expected net payoff that the agent obtains when he undertakes the relationship-specific investment.[11] Following the same logic as in the case of randomization over π_H, if this opportunity cost differs between types, then this may provide a possible rationale for randomization over ϕ_H to be optimal. However, if (**C1**) binds, then the net payoff that the agent obtains when he undertakes the relationship-specific investment under contract M_H is zero for both θ_L and θ_H.[12] Hence, the opportunity cost of randomization over ϕ_H is equal (and null) for both types. Randomization over ϕ_H is therefore ineffective for screening between types.

5 Appendix

5.1 Proofs

Proof of Lemma 1: We show that any menu of contracts in which $\phi_H = 1$ and (2) holds is necessarily dominated, as P could offer a menu that, whilst violating (2), satisfies both IC_H and IC_L and yields him a strictly higher expected payoff. Consider a menu $M = \{M_H, M_L\} = \{(\phi_H, \pi_H, q_H, u_H), (\phi_L, \pi_L, q_L, u_L)\}$ such that $\phi_H = 1$ and (2) holds. P's expected payoff from M is,

$$\lambda \{\pi_H [q_H(v-\theta_H) - C_H] + C_H - B_H - u_H\} + (1-\lambda)\phi_L \{\pi_L [q_L(v-\theta_L) - C_L] + C_L - B_L - u_L\}. \quad (7)$$

Now consider an alternative menu $\widehat{M} = \left\{\widehat{M}_H, \widehat{M}_L\right\}$, where $\widehat{M}_H = (1, \widehat{\pi}_H, \widehat{q}_H, 0)$ and $\widehat{M}_L = (1, 1, \overline{q}, 0)$. Under A1, $\widehat{M}$ satisfies IC_H. It also satisfies IC_L provided,

$$\widehat{\pi}_H \widehat{q}_H \Delta\theta - (1-\widehat{\pi}_H)\Delta C + \Delta B \leq 0 \quad (8)$$

We now show that there exist values of $\widehat{\pi}_H$ and $\widehat{q}_H$ which satisfy (8) with equality (i.e., violate (2)) and which are such that $\widehat{M}$ yields P a greater expected payoff than M. P's expected payoff from $\widehat{M}$ is,

$$\lambda \{\widehat{\pi}_H [\widehat{q}_H(v-\theta_H) - C_H] + C_H - B_H\} + (1-\lambda)[\overline{q}(v-\theta_L) - B_L]. \quad (9)$$

A sufficient condition for (9) to exceed (7) is,

$$\widehat{\pi}_H [\widehat{q}_H(v-\theta_H) - C_H] - \pi_H [q_H(v-\theta_H) - C_H] > 0. \quad (10)$$

[10] It is straightforward to see that in our framework randomization over ϕ_L is also never optimal.

[11] Recall that in our framework the relationship-specific investment is necessary for trade between the principal and the agent to occur. Given a contract $\{\phi, \pi, q, T\}$ if the agent undertakes the relationship-specific investment, his expected utility is $T - \theta_i \pi q + (1-\pi)C_i$, while if the agent does not undertake the relationship-specific investment, his utility is $T + B_i$.

[12] This follows since, when (**C1**) binds, then $-\pi_H q_H \theta_L + (1-\pi_H)C_L - B_L = -\pi_H q_H \theta_H + (1-\pi_H)C_H - B_H = 0$.

Condition (10) ensures that P prefers $\widehat{M}$ to M. We distinguish between two cases. First, suppose that $\frac{(1-\pi_H)\Delta C-\Delta B}{\pi_H\Delta\theta} \leq \overline{q}$. Hence setting $\widehat{\pi}_H = \pi_H$ and $\widehat{q}_H = \frac{(1-\pi_H)\Delta C-\Delta B}{\pi_H\Delta\theta}$ ensures (8) holds with equality. Contract $\widehat{M}_H = \left(1, \pi_H, \frac{(1-\pi_H)\Delta C-\Delta B}{\pi_H\Delta\theta}, 0\right)$ is feasible because, if (2) holds, then $q_H < \frac{(1-\pi_H)\Delta C-\Delta B}{\pi_H\Delta\theta}$ which implies $(1-\pi_H)\Delta C - \Delta B > 0$. With $\widehat{\pi}_H = \pi_H$ the LHS of (10) is $\pi_H\left(\widehat{q}_H - q_H\right)(v-\theta_H)$, which is strictly positive. Hence, $\widehat{M}_H$ dominates M_H and so $\widehat{M}$ dominates M.

Second, suppose $\frac{(1-\pi_H)\Delta C-\Delta B}{\pi_H\Delta\theta} > \overline{q}$. Note that since $\overline{q}\Delta\theta+\Delta B > 0$ under A1, $-\frac{\Delta B}{\Delta\theta} < \overline{q} < \frac{(1-\pi_H)\Delta C-\Delta B}{\pi_H\Delta\theta}$, so $\Delta C - \Delta B > 0$. There are then two possibilities to consider.

In the first case, $q_H\Delta\theta + \Delta C > 0$. Inequality (2) can be rewritten as $\pi_H < \frac{\Delta C-\Delta B}{q_H\Delta\theta+\Delta C}$. By setting $\widehat{\pi}_H = \frac{\Delta C-\Delta B}{q_H\Delta\theta+\Delta C}$, $\widehat{q}_H = q_H$ we ensure (8) holds with equality. The LHS of (10) becomes $(\widehat{\pi}_H - \pi_H)\left[q_H(v-\theta_H) - C_H\right]$, which is strictly positive. Hence, $\widehat{M}_H = \left(1, \frac{\Delta C-\Delta B}{q_H\Delta\theta+\Delta C}, q_H, 0\right)$ dominates M_H and so $\widehat{M}$ dominates M.

In the second case, $q_H\Delta\theta + \Delta C \leq 0$. For this to hold, we require $\Delta C < 0$. As $\Delta C - \Delta B > 0$, this implies $\Delta B < 0$. By setting $\widehat{\pi}_H = 1$, $\widehat{q}_H = -\frac{\Delta B}{\Delta\theta}$ we ensure (8) holds with equality. The LHS of (10) becomes $\left[-\frac{\Delta B}{\Delta\theta}(v-\theta_H) - C_H\right] - \pi_H\left[q_H(v-\theta_H) - C_H\right]$. Under (2), a sufficient condition for this to be positive is that,

$$C_H(v-\theta_L) - C_L(v-\theta_H) < 0. \tag{11}$$

Note however that as $q_H\Delta\theta + \Delta C \leq 0$ in this second case, if (11) does not hold then contract M_H is dominated by a contract that sets $\phi_H = 0$. To see this, note that, by setting $\phi_H = 1$, the extra profit obtained by the principal is non-negative only if $q_H \geq \frac{u_H+B_H-C_H(1-\pi_H)}{(v-\theta_H)\pi_H}$. For this to be consistent with $q_H\Delta\theta + \Delta C \leq 0$ it is necessarily required that $\frac{B_H-C_H(1-\pi_H)}{(v-\theta_H)\pi_H} \leq -\frac{\Delta C}{\Delta\theta}$. In turn, this requires $C_H(v-\theta_L) - C_L(v-\theta_H) < 0$. We therefore conclude that contract M is surely dominated.■

Proof of Proposition 1: We divide the proof in two parts. We first consider the case where (**C1**) is slack in the optimal contract. We then consider the case where (**C1**) binds in the optimal contract. First consider the solution of (**P**) ignoring (**C1**). It is straightforward to see the optimal M_L prescribes $\phi_L = \pi_L = 1$, $q_L = \overline{q}$. Consider now the optimal M_H.

Lemma 2: *If condition* (**C1**) *is slack in the optimal contract, then the optimal* M_H *prescribes* $\pi_H = \phi_H = 1$, $q_H = \overline{q}$.

Proof of Lemma 2: Ignoring (**C1**), the FOCs for M_H are,

$$\frac{\partial U_P}{\partial \pi_H} = \phi_H q_H\left[\lambda(v-\theta_H) - (1-\lambda)\Delta\theta\right] - \phi_H\left[\lambda C_H + (1-\lambda)\Delta C\right] \tag{12}$$

$$\frac{\partial U_P}{\partial q_H} = \phi_H \pi_H\left[\lambda(v-\theta_H) - (1-\lambda)\Delta\theta\right] \tag{13}$$

$$\frac{\partial U_P}{\partial \phi_H} = \lambda\left[\pi_H q_H(v-\theta_H) + (1-\pi_H)C_H - B_H\right] - (1-\lambda)\left[\pi_H q_H\Delta\theta - (1-\pi_H)\Delta C + \Delta B\right] \tag{14}$$

For condition (**C1**) to be slack, it is necessary that $\phi_H = 1$. Hence, the LHS of (14) must be positive.[13] This has implications for the optimal values of π_H and q_H. If the optimal π_H in the unconstrained problem is zero, to then have $\phi_H = 1$ requires $\lambda(C_H - B_H) - (1-\lambda)(\Delta B - \Delta C) > 0$. Since $C_H - B_H < 0$, a necessary condition for this is $\Delta B - \Delta C < 0$. However, when $\phi_H = 1$ and $\pi_H = 0$, this requirement would contradict (**C1**). Similarly, if the optimal q_H in the unconstrained problem is zero, to then have $\phi_H = 1$ requires $\lambda\left[(1-\pi_H)C_H - B_H\right] - (1-\lambda)\left[\Delta B - (1-\pi_H)\Delta C\right] > 0$. For this to hold it is necessary that $\Delta B - (1-\pi_H)\Delta C < 0$. However, when $\phi_H = 1$ and $q_H = 0$, this requirement would again contradict (**C1**). Hence, if (**C1**) is slack in the optimal contract, then the optimal π_H and the optimal q_H must

[13] Recall that, as mentioned in footnote 7, if indifferent between $\phi_H = 1$ and $\phi_H = 0$, P will select $\phi_H = 0$.

both be strictly positive. As mentioned earlier, we adopt the convention that, if indifferent between all possible values of $q_H \in [0, \overline{q}]$, the principal will select $q_H = 0$, and, similarly, if indifferent between all possible values of $\pi_H \in [0,1]$, the principal will select $\pi_H = 0$. Since, as proved above, $q_H = 0$ and/or $\pi_H = 0$ are not consistent with $\phi_H = 1$, we conclude that if (**C1**) is slack in the optimal contract, then the optimal q_H must be $= \overline{q}$, and the optimal π_H must be $= 1$.■

When $\pi_H = \phi_H = 1$, $q_H = \overline{q}$, (**C1**) becomes $\overline{q}\Delta\theta + \Delta B \geq 0$, which is satisfied with strict inequality by A1. It remains to verify consistency; given $\pi_H = \phi_H = 1$, $q_H = \overline{q}$, the LHS of (12), (13) and (14) must be strictly positive, to ensure that $\pi_H = \phi_H = 1$, $q_H = \overline{q}$ is indeed optimal. It is straightforward to show that this happens whenever,

$$\lambda > \max\left\{\frac{\Delta C + \overline{q}\Delta\theta}{\overline{q}\left(v-\theta_L\right) - C_L}, \frac{\Delta\theta}{v-\theta_L}, \frac{\Delta B + \overline{q}\Delta\theta}{\overline{q}\left(v-\theta_L\right) - B_L}\right\}. \tag{4}$$

Hence, when (4) holds, the optimal contracts prescribe $\pi_i = \phi_i = 1$, $q_i = \overline{q}$ for both $i = L, H$. Moreover, $u_H = 0$, while $u_L = \overline{q}\Delta\theta + \Delta B$ (it is straightforward to check that this satisfies IC_H). Condition (**C1**) is slack. This establishes the first part of the proof of proposition 1.

Second part: When (**C1**) binds, $q_H = \frac{(1-\pi_H)\Delta C - \Delta B}{\pi_H \Delta\theta}$, and P's expected payoff is,[14]

$$\begin{aligned} U_P \quad &= \quad \lambda\phi_H\left[\pi_H\left(\frac{(1-\pi_H)\Delta C - \Delta B}{\pi_H\Delta\theta}(v-\theta_H) - C_H\right) + C_H - B_H\right] + \\ &\quad (1-\lambda)\phi_L\left[\pi_L q_L(v-\theta_L) + (1-\pi_L)C_L - B_L\right]. \end{aligned} \tag{15}$$

It is straightforward to see the optimal M_L in this case also prescribes $\phi_L = \pi_L = 1$, $q_L = \overline{q}$. The optimal M_H maximizes (15) subject to $q_H \in [0, \overline{q}]$. The FOCs are,

$$\frac{\partial U_P}{\partial \pi_H} \quad = \quad \lambda\phi_H\left(\frac{-\Delta C(v-\theta_H)}{\Delta\theta} - C_H\right) = \lambda\phi_H\left[C_L(v-\theta_H) - C_H(v-\theta_L)\right] \tag{16}$$

$$\frac{\partial U_P}{\partial \phi_H} \quad = \quad \lambda\left[\pi_H\left(\frac{(1-\pi_H)\Delta C - \Delta B}{\pi_H\Delta\theta}(v-\theta_H) - C_H\right) + C_H - B_H\right] \tag{17}$$

Two cases can arise.[15] In the first $C_L(v-\theta_H) - C_H(v-\theta_L) < 0$, so conditional on $\phi_H = 1$, $\frac{\partial U_P}{\partial \pi_H} < 0$ and P sets π_H as low as possible. If $\Delta C > \Delta B$ then $\frac{\partial q_H}{\partial \pi_H} < 0$ and the lowest feasible π_H solves $\overline{q} = \frac{(1-\pi_H)\Delta C - \Delta B}{\pi_H\Delta\theta}$, so $\pi_H = \frac{\Delta C - \Delta B}{\overline{q}\Delta\theta + \Delta C}$. Provided $\frac{\Delta C - \Delta B}{\overline{q}\Delta\theta + \Delta C}\left(\overline{q}(v-\theta_H) - C_H\right) + C_H - B_H > 0$, the optimal ϕ_H is 1. If $\Delta C < \Delta B$ then $\frac{\partial q_H}{\partial \pi_H} > 0$ and the lowest feasible π_H solves $q_H = 0$. Similarly, if $\Delta C = \Delta B$ then when $\phi_H = 1$ (**C1**) may only bind if $q_H = 0$. However, $\phi_H = 0$ is preferred by P in this case, since it allows P to avoid having to pay a positive transfer (namely, $B_H - (1-\pi_H)C_H$) to type θ_H to induce his participation.

In the second case, $C_L(v-\theta_H) - C_H(v-\theta_L) > 0$, so conditional on $\phi_H = 1$, $\frac{\partial U_P}{\partial \pi_H} > 0$ and P sets π_H as high as possible. If $\Delta B < 0$, then $\pi_H = 1$ and $q_H = -\frac{\Delta B}{\Delta\theta}$. Provided $-\frac{\Delta B}{\Delta\theta}(v-\theta_H) - B_H > 0$, it is then optimal to set $\phi_H = 1$. If $\Delta B > 0$, it is then optimal to set $\phi_H = 0$ as this is the only way to ensure (**C1**) binds. To see this, note that we can only be in the case $C_L(v-\theta_H) - C_H(v-\theta_L) > 0$ if $\Delta C < 0$ so that, if $\Delta B > 0$, then $\Delta C < \Delta B$. This implies $\frac{(1-\pi_H)\Delta C - \Delta B}{\pi_H\Delta\theta} < 0$ for all π_H, and therefore (**C1**) never binds unless $\phi_H = 0$. Similarly, if $\Delta B = 0$, then when $\phi_H = 1$ (**C1**) may only bind if $q_H = 0$. However,

[14] More precisely, $q_H = \frac{(1-\pi_H)\Delta C - \Delta B}{\pi_H\Delta\theta}$ must necessarily hold if $\phi_H = 1$ and (**C1**) binds. If $\phi_H = 0$, then clearly the value of q_H is irrelevant. The expression (15) captures P's expected payoff in both cases.

[15] The knife-edge case where $C_L(v-\theta_H) - C_H(v-\theta_L) = 0$ is ignored for brevity.

as argued above, $\phi_H = 0$ is then preferred by P. To complete the description of the optimal contracts, note that when (**C1**) binds the optimal contracts prescribe $u_H = u_L = 0$. It is straightforward to check IC_H is satisfied in all the cases we have identified.■

5.2 Assumption A1 Does Not Hold

For completeness, we consider the case in which $0 \geq \overline{q}\Delta\theta + \Delta B$ and so θ_H types have incentives to understate their costs and mimic θ_L types. The remaining assumption A2 is assumed to still hold. The counterparts for the main results are as follows,

Lemma 1B: *It is never optimal for the principal to offer $\phi_L = 1$ in conjunction with π_L and q_L satisfying,*

$$-\pi_L q_L \Delta\theta + (1 - \pi_L)\Delta C - \Delta B < 0. \tag{18}$$

An implication is that the participation constraint of type θ_H will not bind at the optimum. The optimal contracts are now found by letting IC_H hold with equality, setting $u_L = 0$, and ignoring IC_L. The counterpart to (**C1**) is,

$$\phi_L \left[-\pi_L q_L \Delta\theta + (1 - \pi_L)\Delta C - \Delta B\right] \geq 0. \tag{C1B}$$

Proposition 2B: *For type θ_H, the optimal contract always prescribes $\phi_H = \pi_H = 1$, $q_H = \overline{q}$. If*

$$\lambda < \min\left\{\frac{\overline{q}\,(v - \theta_L) - C_L}{\overline{q}\,(v - \theta_H) - C_H}, \frac{v - \theta_L}{v - \theta_H}, \frac{\overline{q}\,(v - \theta_L) - B_L}{\overline{q}\,(v - \theta_H) - B_H}\right\} \tag{19}$$

*then (**C1B**) is slack, and the optimal contract for type θ_L has $\phi_L = \pi_L = 1$, $q_L = \overline{q}$. If* (19) *doesn't hold, then* (**C1B**) *binds, and the optimal contract for type θ_L is,*

(i) if $C_H < \frac{C_L(v-\theta_H)}{v-\theta_L}$ and $\Delta C - \Delta B < \frac{(B_L - C_L)(\overline{q}\Delta\theta + \Delta C)}{\overline{q}(v-\theta_L) - C_L} < 0$: $\phi_L = 1$, $\pi_L = \frac{\Delta C - \Delta B}{\overline{q}\Delta\theta + \Delta C}$ and $q_L = \overline{q}$.

(ii) if $C_H > \frac{C_L(v-\theta_H)}{v-\theta_L}$ and $\Delta B < -\frac{B_H\Delta\theta}{v-\theta_L} < 0$: $\phi_L = \pi_L = 1$ and $q_L = -\frac{\Delta B}{\Delta\theta}$.

(iii) in all the other cases: $\phi_L = 0$.

5.3 A Numerical Example

Suppose $\theta_H = 0.75$, $\theta_L = 0.25$, $\overline{q} = v = 2$, and agent's *ex ante* and *ex post* outside options are $B_H = 1.85$, $B_L = 2.35$, $C_H = 1.75$, and $C_L = 1.95$. For (4) to hold we require $\lambda \geq 0.52$. If $\lambda < 0.52$, then (**C1**) must bind in the optimal contract. From (3), if $\phi_H = 1$ this implies $q_H = \frac{2}{5} + \frac{3}{5\pi_H}$, and to ensure $q_H \leq \overline{q} = 2$, we require $\pi_H \geq 0.375$. Conditional on $\phi_H = 1$, the principal then selects $\pi_H \in [0.375, 1]$ to maximize her expected payoff when dealing with a type θ_H agent, $U_P = \pi_H \left[\left(2/5 + \frac{3}{5\pi_H}\right) 1.25 - 1.75\right] - 0.1$. Since U_P is decreasing in π_H, so the principal selects the lowest π_H compatible with (**C1**). The optimal contract for θ_H then is, $\phi_H = 1, \pi_H = 0.375, q_H = \overline{q} = 2$, and when dealing with type θ_H agents, the principal's expected payoff is 0.18.

References

[1] ARNOTT.R AND J.E.STIGLITZ (1988) "Randomization with Asymmetric Information", *Rand Journal of Economics* 19: 344-62.

[2] BARON.D.P AND R.B.MYERSON (1982) "Regulating a Monopolist with Unknown Costs", *Econometrica* 50: 911-30.

[3] BRITO.D.L, J.H.HAMILTON, S.M.SLUTZKY, AND J.E.STIGLITZ (1995) "Randomization in Optimal Income Tax Schedules", *Journal of Public Economics* 56: 189-223.

[4] CALZOLARI.G. AND A.PAVAN (2006) "On the Optimality of Privacy in Sequential Contracting", *Journal of Economic Theory* 130: 168-204.

[5] GROSSMAN.S.J AND O.D.HART (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration", *Journal of Political Economy* 94: 691-719.

[6] GROUT.P A (1984) "Investment and Wages in the Absence of Binding Contracts: a Nash Bargaining Approach", *Econometrica* 52: 449-60.

[7] GUESNERIE.R AND J-J.LAFFONT (1984) "A Complete Solution to a Class of Principal-Agent Problems With an Application to the Control of a Self-Managed Firm", *Journal of Public Economics* 25: 329-69.

[8] JULLIEN.B (2000) "Participation Constraints in Adverse Selection Models", *Journal of Economic Theory* 93: 1-47.

[9] HART.O.D AND J.MOORE (1990) "Property Rights and the Nature of the Firm", *Journal of Political Economy* 98: 1119-58.

[10] LEWIS.T AND D.SAPPINGTON (1989) "Countervailing Incentives in Agency Problems", *Journal of Economic Theory* 49: 294-313.

[11] MAGGI.G AND A.RODRIGUEZ-CLARE (1995) "On Countervailing Incentives", *Journal of Economic Theory* 66: 238-63.

[12] MIRRLEES.J.A (1971) "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies* 38: 175-208.

[13] MYERSON.R (1981) "Optimal Auction Design", *Mathematics of Operations Research* 6: 58-73.

[14] ROCHET.J-C (1984) Monopoly Regulation with Two-Dimensional Uncertainty, mimeo Université Paris.

[15] STIGLITZ.J.E (1982) "Self-Selection and Pareto Efficient Taxation", *Journal of Public Economics* 17: 213-40.

[16] STRAUSZ.R (2006) "Deterministic Versus Stochastic Mechanisms in Principal-Agent Models", *Journal of Economic Theory* 128: 306-14.

[17] THANASSOULIS.J. (2004) "Haggling Over Substitutes", *Journal of Economic Theory* 117: 217-245.