Moral Hazard, Optimal Contracting and Strategic Competition

David Deller

Current Draft: June 2012

Abstract

This paper opens up the "black box" of firms' profit functions to consider the affect of modelling firms as principal-agent relationships on the outcomes of strategic competition. Using a parameterised model it is found that for large market sizes moral hazard reduces expected output more severely than if firms restricted output to jointly maximise profits i.e. collude. This result is obtained even though firms use fully optimal non-linear incentive contracts. Additionally profit-maximising firms may not always invest in a perfect monitoring technology even if it is welfare enhancing. As such there may be an over-reliance on incentive contracts from a welfare perspective. Overall this paper highlights the importance of modelling firms' internal workings when addressing regulatory and other industrial organisation questions.

Acknowledgements: I would like to thank particularly Prof. Francesco Squintani and Prof. Melvyn Coles for their advice and encouragement with this paper. For their valuable additional input I would like to thank Dr. Kate Rockett, Prof. Ken Burdett, Prof. Pierre Regibeau, fellow research students at the University of Essex and attendees at Royal Holloway's Annual PhD Conference 2012, the NIE's Doctoral Student Colloquium 2011, the ESRC Centre for Competition Policy's summer conference 2010 and the RES's Autumn School 2010. This research was completed whilst in receipt of a "1+3" studentship from the ESRC.

1 Introduction

This paper's purpose is to highlight the impact a classic moral hazard problem within firms can have on equilibrium outcomes in a market involving strategic competition between firms. The paper demonstrates that even when companies are able to write fully optimal contracts the costs of agency can still have a significant downward impact on equilibrium output. Indeed, in a parameterised model with linear demand, as markets grow "large" the downward impact of moral hazard on expected output is greater than if firms could collude to restrict output. The significance of this result is that the welfare losses associated with collusion are seen as justifying the policy responses of antitrust authorities. The key question, therefore, is whether there is a role for policy regarding agency costs?

Some level of agency costs will often be inherent in a production process. It is shown that profit-maximising firms may fail to make investments that reduce agency costs despite them being welfare enhancing. As such, from a welfare perspective, there maybe an over-reliance on incentive contracts. The remaining open question is whether policymakers/regulators can actually make welfare enhancing interventions.

Overall this paper highlights the importance of modelling firms as collections of utility maximising individuals, bound together by contracts, when considering firm behaviour in oligopolies. It also shows why consumers and policymakers, not just shareholders, have a legitimate interest in the way firms resolve agency problems.

A myriad of different agency problems face firms and their seriousness, particularly for large corporations, has long been recognised.[1] This paper focuses on perhaps the classic agency problem: unobservable effort by a risk-averse agent.[2] In the current model agent effort alters the probability distribution of a firm's output with higher effort increasing the likelihood of higher output. Each firm comprises a single principal and a single agent. The agent, therefore, is best interpreted as a senior production manager whose effort alters an entire firm's output.

Since the agent selects their effort based on the incentive contract offered the principal can alter the firm's expected output by altering the incentive contract. To the best of my knowledge this is the first model of agency costs in oligopoly where a cost-minimising non-linear incentive contract, incorporating effort and a performance measure as continuous variables, is derived using the first-order approach. In a generalised n-firm setting conditions are obtained for the existence of an equilibrium in the principals' contract parameter choice game. In a parameterised setting the model is solved with both linear and exponential inverse demand curves. Using a linear inverse demand curve the results regarding market size are obtained. Using an exponential inverse demand curve allows analysis regarding the number of firms and the elasticity of demand to be performed.

A short literature review follows in section 2. In section 3 a parameterised version of the model is introduced and in section 4 this model is solved. Section 5 analyses the model numerically including for several extensions. Section 6 provides a discussion and considers the results' robustness before section 7 concludes.

[1] See Jensen and Meckling (1976).
[2] See Stiglitz (1974).

2 Literature Review

Starting with Leibenstein's (1966) discussion of "X-inefficiency", it has been widely theorised that competition between firms can alter the extent of agency problems in firms. However research investigating the impact of agency problems on market outcomes is more limited. This is despite the prevalence of agency relationships within companies, the considerable attention given by researchers and policymakers to competition and the continuing debates about executive pay. Examples of work looking at the impact of competition on agency problems include Hart (1983), Scharfstein (1988) and Schmidt (1997). Hart (1983) and Scharfstein (1988) both analyse the impact a fringe of "entrepreneurial" firms without agency problems has on firms that have a separation of ownership and control. Schmidt (1997) emphasises that greater competition has a disciplining effect on employees by increasing instead the threat of liquidation and unemployment. However, when taken as a whole, this literature provides mixed results regarding the impact of increased competition on effort incentives.

Closer to the current paper is Raith (2003). Raith constructs a circular city model of monopolistic competition which allows entry by firms. Here agent effort results in a reduction in marginal costs. However Raith restricts attention to linear incentive contracts and the paper's aim is different to the current investigation. Raith's purpose is to create a setting where there is a positive relationship between risk and incentive strength.[3]

The strategic delegation literature, started by Fershtman and Judd (1987) and Sklivas (1987), does investigate the impact of incentive contracts on product market competition. However, whilst this literature is described in terms of a "principal" and an "agent", it generally assumes agents are risk-neutral and so agency costs are ignored. Instead the role of delegation itself as a strategic commitment device is emphasised. Whilst accepting that incentive contracts can be used strategically the current paper takes the original theoretical basis for incentive contracts - to induce effort - as the primary motivation for their existence.

Gal-Or (1997) provides an overview of the small number of contributions to the strategic delegation literature that do include true agency problems. However in these papers, such as Fumas (1992) and Gal-Or (1993), attention is again restricted to linear incentive contracts or the focus remains on decisions regarding delegation and organisational structure. The salesforce compensation literature, for example Bhardwaj (2001), also looks at delegation questions whilst incorporating moral hazard problems. In this literature Mishra and Prasad (2005), unusually, do consider optimal

[3] Other papers looking at the impact of competition on incentives include Holden (2008), Plehn-Dujowich and Serfes (2010) and Theilen (2010).

non-linear incentive contracts. However Mishra and Prasad only demonstrate the combinations of centralised pricing and delegation that can be equilibrium candidates.

Moving away from delegation issues Hermalin (1994) demonstrates that non-convexities in firms' agency problems can cause otherwise identical firms to offer incentives of differing strengths. While Aghion, Dewatripont and Rey (1999) consider the role of external financing as a disciplining device on managers of firms engaged in oligopolistic competition. Lastly Bonatti (2003) suggests that the effort monitoring capabilities of unions may favour the use of collective bargaining in oligopoly settings.

3 The Model

This section describes the parameterised model's structure. The model is presented from the perspective of principal-agent pair i.

Two principal-agent pairs (firms) compete in a quantity competition game. The overall game involves three stages:

Stage 1 - Simultaneously each principal devises the optimal incentive contract for their agent. This includes selecting the optimal effort level to induce.

Stage 2 - Given the incentive contract offered agents select the effort level maximising their utility.

Stage 3 - The outcome of the production process is realised. The output is sold at the price required to clear the market. All players receive their pay-offs.

The principal-agent pairs are assumed identical in all respects.[4] Output is homogeneous and the output of each firm is a random variable dependent on agent effort. Firm i's output is denoted q_i where $q_i \in [0, \infty)$ and $Q = q_i + q_j$. Agent i's effort is denoted a_i where $a_i \in [\underline{a}, \overline{a}]$, $\underline{a} > 0$ and $A = a_i + a_j$.[5] The inverse demand function is linear:

$$P(Q) = \left\{ \begin{array}{cc} B - Q, & B \geq Q \\ 0, & B < Q \end{array} \right\}$$

Output is exponentially distributed with the probability density function for output given a specific effort level a_i being:

$$f(q_i|a_i) = \left\{ \begin{array}{cc} \frac{1}{a_i} e^{-\frac{q_i}{a_i}}, & q_i \geq 0 \\ 0, & q_i < 0 \end{array} \right\}$$

[4] The proof for existence of equilibrium, shown in Appendix 1, does not require symmetry.

[5] Assume that $\underline{a}$ is sufficiently low and $\overline{a}$ is sufficiently high that they never impinge on the equilibrium outcome.

The exponential distribution is used for several reasons. Firstly, this distribution's support is $[0, \infty)$ thereby ruling out negative quantities. Secondly, since the distribution's support remains constant regardless of effort the principal can never use output to infer the an agent's effort with certainty. Thirdly, it ensures Jewitt's (1988) conditions for the first-order approach to be valid hold and, lastly, the cumulative distribution for output, $F(q_i|a_i)$, satisfies the condition for the single-crossing property to hold. The condition for single-crossing is that $F_{a_i}(q_i|a_i) \leq 0$ holds for all q_i and for some q_i $F_{a_i}(q_i \mid a_i) < 0$ holds. As such output is an informative signal of effort.

Each principal is risk-neutral and principal i's utility function is $Y(\pi_i) = \pi_i$ where π_i is the principal's expected pay-off. Agent i's utility is denoted $U(w_i)$ where w_i is agent i's income.[6] Since the agent is risk-averse $U(w_i)$ is increasing concave in w_i. The agent's total utility is denoted $\Omega\left(w_{i,}a_i\right)$ and the disutility from effort is denoted $V(a_i)$. The agent's utility from income and disutility from effort are additively separable so that $\Omega\left(w_{i,}a_i\right) = U(w_i) - V(a_i)$. In the model solved below $V(a_i) = a_i^2$ and $U(w_i) = 2\left(w_i\right)^{\frac{1}{2}}$.[7]

As effort is unobservable assume that the principal must use an incentive compatible contract that is a continuous function of output to induce effort. Denote this contract $w_i(q_i)$. As the distributions of output for each firm are independent no reduction in costs is offered by using relative performance evaluation and rewarding agents on the basis of both q_i and q_j. Similar reasoning also means there is no benefit from rewarding an agent on the basis of their firm's profits.

The labour market is competitive and denote the agent's reservation utility $R \geq 0$. Without a loss of generality assume there are no other production costs beyond the cost of labour. Also assume that each principal must always operate, employ an agent and induce the minimum effort level $\underline{a}$.[8]

Beyond the effort exerted there is no other hidden information in the model. It is assumed that each principal knows the shape of their own agent's utility function and all the characteristics of the rival principal-agent pair.

4 Solving the model

The objective when solving the model is to obtain a pure strategy Nash equilibrium for the contract parameters, also denoted a_i (a_j), selected by each principal.

[6] For simplicity assume the agent has no other wealth or income sources.

[7] This particular utility function is used for tractibility.

[8] This assumption ensures no downward jump in the firm's reaction function occurs which could affect the proof of existence. As such the model considers a short-run setting.

Principal i's problem can be thought of as involving two separate steps. The first is to derive a contract which induces a given level of effort from agent i at the minimum cost. The second is to select the amount of effort to induce, via the incentive contract, to maximise profits given the interaction between firms in the product market. Each of these steps is considered in turn.

Step 1: Deriving the cost-minimising contract:

The incentive contract must satisfy the agent's participation constraint (PC) and incentive compatibility constraint (IC). Formally this problem can be expressed as:

$$\max_{w_i(q_i)} \int_0^\infty -w_i(q_i) dF(q_i|a_i) \qquad (1)$$

subject to PC:

$$\int_0^\infty 2 \left(w_i(q_i)\right)^{\frac{1}{2}} dF(q_i|a_i) - a_i^2 \geq R \qquad (2)$$

and IC:

$$\int_0^\infty 2 \left(w_i(q_i)\right)^{\frac{1}{2}} dF_{a_i}(q_i|a_i) - 2a_i = 0 \qquad (3)$$

The cost-minimising contract is found using the first-order approach as in Mirrlees (1976) and Holmstrom (1979). To be incentive compatible the contract must maximise the agent's expected utility at the effort level the principal wishes to induce. The first-order approach uses the first-order condition (FOC) of the agent's maximisation problem as a sufficient condition for reaching the global maximum of the agent's problem. Thus the agent's FOC is used as the IC in the principal's maximisation problem above.

Mirrlees (1999)[9] identified that using an agent's FOC as the IC is not generally valid. The FOC is a necessary rather than sufficient condition for maximising the agent's utility. However Jewitt (1988) provides conditions for which the FOC is a sufficient condition for maximisation of an agent's utility. Jewitt's conditions on the distribution function and utility function are:

(i) $\int_{-\infty}^{q} F(q|a)dq$ is non-increasing convex in a for each value of q

(ii) $\int qdF(q|a)dq$ is non-decreasing concave in a

(iii) $\frac{f_a(q|a)}{f(q|a)}$ is non-decreasing concave in q for each value of a

(iv) the utility of the agent is a concave increasing function of the observable variables; mathematically $\omega(z) = U\left(U'^{-1}\left(\frac{1}{z}\right)\right)$, where $z > 0$, is concave

An explanation of why these conditions are needed is provided in Appendix 2.

[9] This paper was originally completed in 1975 but not published.

Lemma 1 *Jewitt's (1988) conditions for the validity of the first-order approach are met by the current parameterised model.*

Proof. Jewitt states that all the distributions falling within the exponential family meet conditions (i)-(iii) and hence the current model satisfies these conditions. Regarding condition (iv), since $\frac{1}{U'(w)} = (w(q))^{\frac{1}{2}}$, $U(w)$ is a linear, and therefore concave, transformation of $\frac{1}{U'(w)}$. Hence using the first-order approach is valid. ■

Whilst Jewitt's conditions ensure the agent's maximisation problem is concave they do not ensure the concavity of the principal's maximisation problem.[10] At present the concavity (strict quasiconcavity) of the principal's problem is assumed.

Lemma 2 *The cost-minimising contract for the principal to induce an effort level* a_i *is:*

$$w_i^*(q_i) = \tfrac{1}{4}\left(a_i^2 + R + 2a_i\left(q_i - a_i\right)\right)^2$$

Proof. Principal i's problem can be expressed as the following Lagrangian :

$$\begin{aligned} \max_{w_i(q_i)} L_i\left(a_i\right) &= \textstyle\int_0^\infty -w_i(q_i)dF(q_i|a_i) \\ +\lambda_i \left[\textstyle\int_0^\infty 2\left(w_i(q_i)\right)^{\frac{1}{2}} dF(q_i|a_i) - a_i^2 - R\right] &+ \mu_i \left[\textstyle\int_0^\infty 2\left(w_i(q_i)\right)^{\frac{1}{2}} dF_{a_i}(q_i|a_i) - 2a_i\right] \end{aligned} \tag{4}$$

In (2) the PC is expressed as an inequality constraint however in (4) it is assumed to bind with equality. Intuitively this assumption must be true. If the PC did not bind with equality the principal could reduce the transfer payment made to the agent, thus increasing profits, whilst still ensuring the agent exerted the desired effort level.

The necessary condition for the cost-minimising contract is found by holding a_i fixed and taking the partial derivative of (4) with respect to $w_i(q_i)$. Setting equal to zero gives:

$$\begin{aligned} \tfrac{\partial L_i}{\partial w_i(q_i)} = &- \textstyle\int_0^\infty dF(q_i|a_i) + \lambda_i \int_0^\infty \left(w_i(q_i)\right)^{-\frac{1}{2}} dF(q_i|a_i) \\ &+\mu_i \textstyle\int_0^\infty \left(w_i(q_i)\right)^{-\frac{1}{2}} dF_{a_i}(q_i|a_i) = 0 \end{aligned}$$

Dividing throughout by $\left(w_i(q_i)\right)^{-\frac{1}{2}} dF(q_i|a_i)$, recognising that $dF(q_i|a_i) = f(q_i|a_i)dq_i$ and re-arranging leads to:

$$\left(w_i(q_i)\right)^{\frac{1}{2}} = \lambda_i + \mu_i \tfrac{f_{a_i}(q_i|a_i)}{f(q_i|a_i)}$$

Inserting the expressions for $f(q_i|a_i)$ and $f_{a_i}(q_i|a_i)$ and then simplifying gives:

[10] The need for the principal's problem to be concave is noted by Grossman and Hart (1983).

$$w_i(q_i) = \left(\lambda_i + \frac{\mu_i}{a_i^2}(q_i - a_i)\right)^2 \qquad (5)$$

Inserting (5) into the IC and PC and then solving as a system of two equations in two unknowns gives the following expressions for λ_i and μ_i:

$$\lambda_i = \frac{a_i^2 + R}{2} \qquad (6)$$
$$\mu_i = a_i^3 \qquad (7)$$

Inserting these values for λ_i and μ_i back into (5) gives the cost-minimising contract to induce the effort level a_i as:

$$w_i^*(q_i) = \frac{1}{4}\left(a_i^2 + R + 2a_i(q_i - a_i)\right)^2 \qquad (8)$$

■

Step 2: Selecting the optimal value of a_i:

The second step of the principal's problem is to select the optimal value of a_i to write in the incentive contract given by (8). Note that a_i is both a parameter in the contract and also the level of effort agent i will exert given (8). As such when selecting the optimal value of a_i to write in the contract the principal is selecting the optimal effort level to induce in the agent.

Holding principal j's choice of a_j fixed the unconstrained profit maximisation problem facing principal i is:

$$\max_{a_i} E(\pi_i) = \int_0^B \int_0^{B-q_j} (B - q_i - q_j)\, q_i dF(q_i|a_i) dF(q_j|a_j) - \int_0^\infty w_i^*(q_i) dF(q_i|a_i)$$

Substituting in the expressions for $w_i^*(q_i)$, $dF(q_i|a_i)$ and $dF(q_j|a_j)$, setting $R = 0$[11] and integrating gives firm i's expected profit function as:

$$E(\pi_i) = a_i a_j^3 \frac{e^{-\frac{B}{a_j}}}{(a_i - a_j)^2} + a_i^2 e^{-\frac{B}{a_i}} \frac{Ba_i - Ba_j + 2a_i^2 - 3a_i a_j}{(a_i - a_j)^2} + (B - 2a_i - a_j)\, a_i - \frac{5}{4} a_i^4 \qquad (9)$$

Assuming that $E(\pi_i)$ is strictly quasiconcave and there exists a point such that $\frac{\partial E(\pi_i)}{\partial a_i} = 0$ then the FOC, $\frac{\partial E(\pi_i)}{\partial a_i} = 0$, will be a necessary and sufficient condition for profit maximisation. The full FOC is:

$$\frac{\partial E(\pi_i)}{\partial a_i} =$$
$$-a_j^3 \frac{e^{-\frac{B}{a_j}}}{(a_i - a_j)^3}(a_i + a_j) + \frac{e^{-\frac{B}{a_i}}}{(a_i - a_j)^3} \begin{pmatrix} B^2 a_i^2 - 2B_i^2 a_i a_j + B^2 a_j^2 + 3Ba_i^3 \\ -8Ba_i^2 a_j + 5Ba_i a_j^2 + 4a_i^4 - 11a_i^3 a_j + 9a_i^2 a_j^2 \end{pmatrix}$$
$$+B - 4a_i - a_j - 5a_i^3 = 0 \qquad (10)$$

[11] This is done for simplicity.

An equivalent condition can be found for principal j.

The next step is to demonstrate an equilibrium exists in the principals' contract parameter choice game.

Theorem 1 *In a n-firm setting if:*
(i) Jewitt's (1988) conditions for the validity of the first-order approach hold, and
(ii) each principal's profit function is strictly quasiconcave in their own contract parameter (a_i for principal i)
then an equilibrium will exist in the principals' contract parameter choice game.

Proof. The proof is adapted from the proof of existence of an n-firm Cournot equilibrium by Frank and Quandt (1963). The proof is shown in Appendix 1.[12]

Given Theorem 1 the equilibrium values of a_i and a_j can be found numerically using (10) and the equivalent condition for principal j.

5 Numerical Analysis

5.1 Varying market size (B)

To provide comparative benchmarks for the impact of moral hazard on market outcomes two additional scenarios are considered. The first is where effort is observable and verifiable thus removing the agency problem. This scenario is referred to as the first best. In the second scenario effort is again observable and verifiable but now firms act to maximise joint profits, i.e. collude.[13] This second benchmark is chosen as the lost output/surplus from collusion is seen as sufficient to justify a policy response of antitrust laws.

[12] If the non-standard demand function:

$$P(q_i, q_j) = B - q_i - q_j \text{ for all } q_i \geq 0, q_j \geq 0$$

an unique equilibrium exists in the principals' contract parameter choice game for the parameterised version of the model. The proof for this is an adaptation of Szidarovsky and Yakowitz's (1977) proof of an unique Cournot equilibrium. Imposing additional assumptions this proof can be extended to a general n-firm setting.

Significantly, in a parameterised model using this non-standard demand function, it is possible to prove that each principal's profit maximisation problem is concave. It is also possible to show that as B grows large the equilibrium values of a_i and a_j found using this non-standard demand function tend to those using the standard demand function stated in section 2. The workings for these results are available on request.

[13] The ability to sustain collusion is assumed rather than demonstrated.

Note that for the exponential distribution the expected output of a firm equals the effort exerted by the firm's agent, i.e. $E(q_i) = a_i$.

To solve the problem numerically it is easiest to appeal to the problem's symmetry and consider a symmetric equilibrium such that $a_i^* = a_j^* = a$. Using this fact, setting $R = 0$ and applying l'hôpital's rule three times to (10) gives the equilibrium condition for moral hazard in the table below. The proofs of the other two conditions are included in Appendix 3.

First Best	$\frac{1}{a^2}e^{-\frac{B}{a}}\left(\frac{1}{3}B^3 + \frac{3}{2}B^2a + 4Ba^2 + 6a^3\right) + B - 5a - a^3 = 0$
Moral Hazard	$\frac{1}{a^2}e^{-\frac{B}{a}}\left(\frac{1}{3}B^3 + \frac{3}{2}B^2a + 4Ba^2 + 6a^3\right) + B - 5a - 5a^3 = 0$ [14]
Maximisation of joint profits	$\frac{1}{2a^2}e^{-\frac{B}{a}}\left(B^3 + 4B^2a + 10Ba^2 + 12a^3\right) + B - 6a - a^3 = 0$

Proposition 1 *As the market becomes "large" moral hazard has a far greater downward impact on expected output per firm than maximisation of joint profits (collusion) by firms.*

This central result is shown in Figure 1. In the top graph expected output per firm is plotted for values of B from 0.05 to 5. This highlights that for "small" market sizes collusion causes a greater reduction in expected output than moral hazard. The bottom graph considers values of B from 0.05 to 250. It illustrates moral hazard's greater downward impact on expected output as the market becomes "large".[15][16]

The relative importance of moral hazard versus collusion varies according to the relative sizes of agency costs and the negative revenue externality associated with competition. The relative importance of moral hazard and collusion changes with market size will therefore depend on the specifications of the demand curve, the cost of effort function and the utility function. The cost of effort function and the utility function together determine the agency cost. The specification of the demand curve determines the size of the negative revenue externality that occurs when firms operate independently.[17] This externality is internalised when firms act together to maximise their joint profits. The larger the externality to be internalised the greater the drop in expected output when firms maximise joint profits.

[14] Given that expected wage costs remain convex when there is observable and verifiable effort Theorem 1 also holds for the first best and maximisation of joint profits.

[15] The MATLAB M-files generating Figures 1, 3, 4 and 5 are available on request.

[16] If the agents' reservation utility, R, is high enough then for all values of B large enough to generate positive expected profits moral hazard will have a greater downward impact on expected output than maximisation of joint profits.

[17] When setting their own contract parameter independent firms will fail to consider the impact their choice has on the revenue received by rival firms. If firm i increases its contract parameter it increases firm i's expected output and hence lowers the expected price for firm j's output.

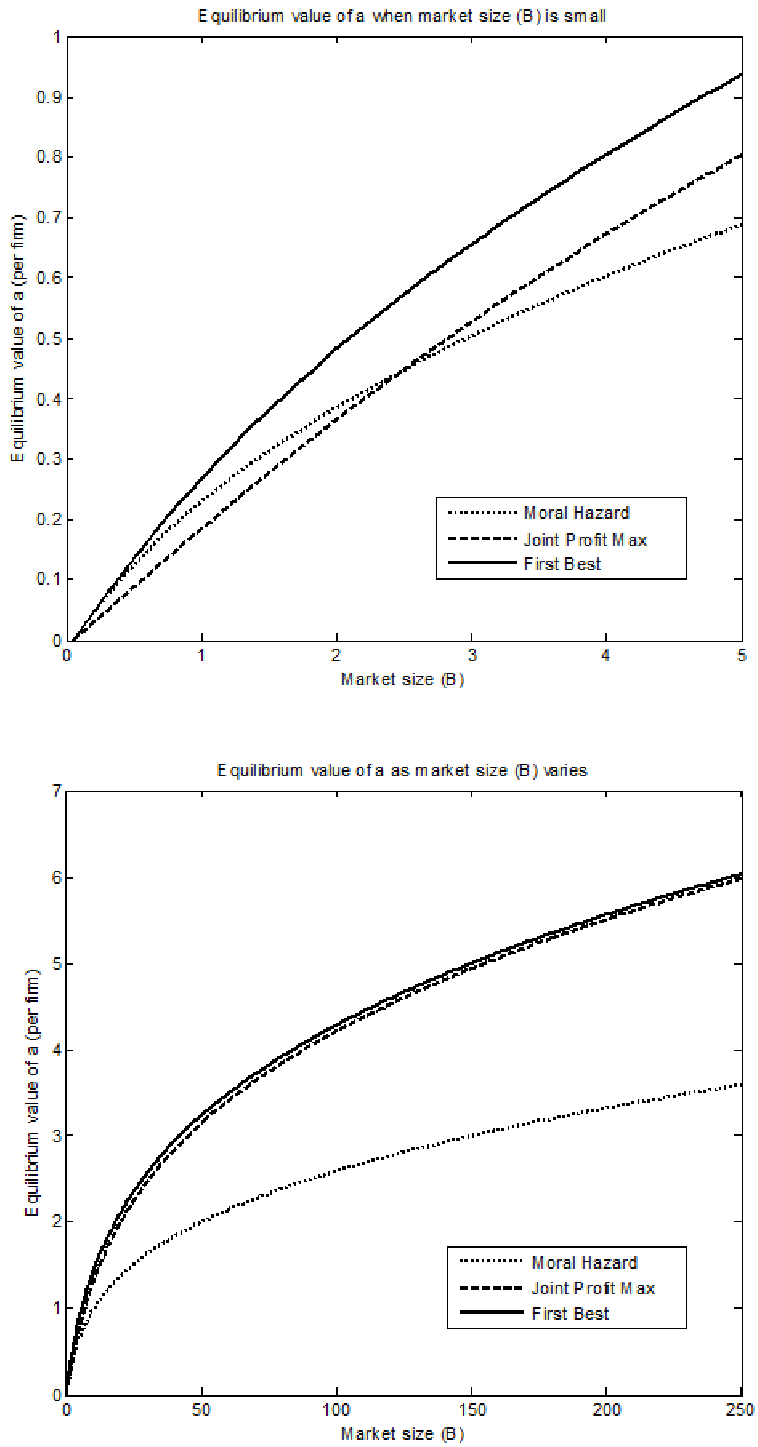


Figure 1: Equilibrium expected output (contract parameter/effort level) per firm as market size, B, increases

In the current setup the agency cost (the difference in wage costs between when moral hazard is and is not present) is a_i^4 and so is convex in effort. Assuming B is large relative to a_i the externality firm i causes is approximately $-a_i a_j$.[18] Therefore when market size increases and firms induce more effort the agency cost increases at a faster rate than the externality. Hence when B is large a firm facing a moral hazard problem will set a lower value of a than one maximising joint profits.

5.2 Extension 1 - Investment in a perfect monitoring technology

A key question is whether the above results have policy implications? There is a potential role for policy only if a benevolent social planner could increase welfare. If agency costs are an inherent part of the production process there will be no role for policy. However it is probable that firms do have means to reduce agency costs. For example, firms could invest in productivity monitoring software or employ "mystery shoppers". The issue is whether the incentives for profit-maximising firms to minimise agency costs lead to welfare maximising outcomes. It is already known, from Shleifer and Vishny (1986), that if the ownership of a firm is dispersed, i.e. there are multiple principals, then there may be an under-investment in monitoring. This is because each principal can free-ride on the monitoring effort of others. This section highlights that under-investment in monitoring may occur for another reason: firms do not consider the gains in consumer surplus associated with reduced agency costs.

For simplicity, assume firms can invest in a perfect monitoring technology that makes effort observable and verifiable. Let each firm have a discrete action set $D \in \{Invest, NotInvest\}$. Simultaneously each firm decides which action to play from D in a stage game prior to the game laid out in section 2. The cost of the investment to each firm is C.

Three numerical examples are considered where B is set to 50, 100 and 200 respectively. The purpose is to see if the range of C where profit-maximising firms choose $Invest$ matches the range of C where a social planner maximising total surplus would play $Invest$.

Proposition 2 *There exists a range of investment costs such that profit-maximising firms will not undertake an investment in a perfect monitoring technology despite it being welfare enhancing. As market size, B, grows larger so does the region of investment costs where differing decisions are made.*

[18] Recall from footnote 14 that when B is large the current model can be approximated with one where the inverse demand curve is: $P(q_i, q_j) = B - q_i - q_j$ for all $q_i, q_j \geq 0$. The value $-a_i a_j$ comes from this alternative model.

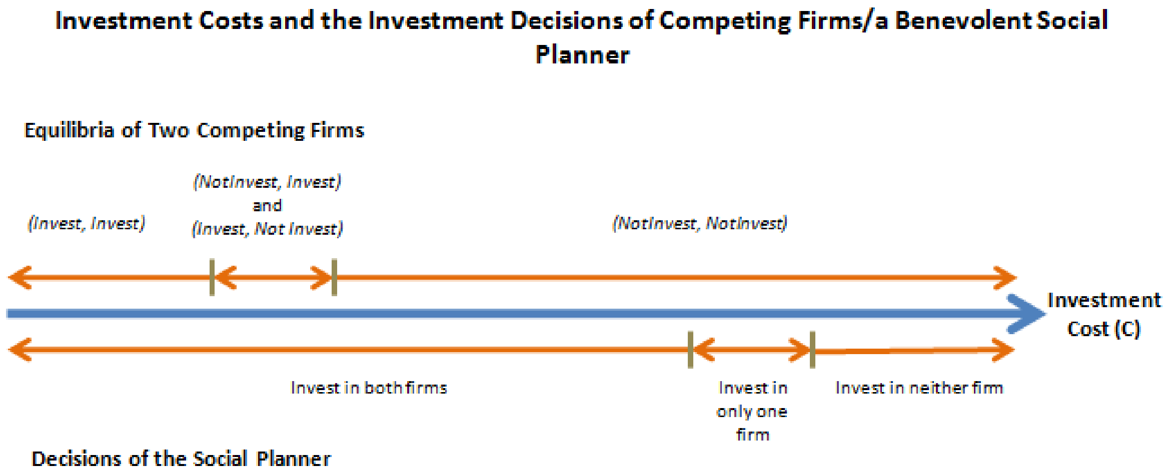


Figure 2: The differing investment decisions of independent firms and a benevolent social planner

Proof: See Appendix 4

The nature of the divergence in investment decisions is summarised in Figure 2. Where there is a divergence in the investment decisions one can say that profit-maximising firms will have an over-reliance on incentive contracts from a welfare perspective.

The size of the region where the investment choices of firms and a social planner differ will depend on the model's specification, in particular, the form of the inverse demand curve.[19] However the underlying intuition for this divergence appears robust. Profit-maximising firms always will be less willing to invest than a welfare maximising social planner unless firms can fully capture their products' consumer surplus. As the market becomes larger consumer surplus increases and so the investment cost range where investment decisions diverge becomes larger.

5.3 Extension 2 - An exponential inverse demand curve

Whilst using a linear inverse demand curve provides consistency with basic oligopoly models the resulting expected revenue function significantly complicates the analysis. The analysis is simplified when, instead, an exponential inverse demand curve is used. Significantly this change of specification makes it possible to prove that the expected profit function is strictly quasiconcave in the parameterised setting. It also allows

[19] See Proposition A4.1 in Appendix 4 for the numerical ranges.

analysis regarding the number of firms in an industry and the elasticity of the inverse demand curve.

The inverse demand curve used in this extension is:

$$P(Q) = De^{-\tau Q} \qquad (11)$$

where D and τ are strictly positive constants and Q is the total output produced by all firms. D represents the highest willingness to pay of any consumer (i.e. the vertical intercept of the inverse demand curve) whilst τ affects the elasticity of the inverse demand curve. The elasticity of price with respect to quantity which is given by:

$$\varepsilon = \frac{dP}{dQ}\frac{Q}{P} = -\tau Q$$

An increase in τ increases this elasticity's absolute value thus making price more sensitive to the volume of output produced.

Theorem 1 can be used to prove that an n-firm equilibrium exists in a generalised setting for the inverse demand curve given by (11). The key assumption remains that the expected profit function is strictly quasiconcave. Apart from changing the expected revenue function in (A1.4) there are no other changes to the proof.

Lemma 3 *For the exponential inverse demand function described in (11) the expected profit function will be strictly quasiconcave in a_i if the assumptions of the parameterised model hold. Hence a n-firm equilibrium in the contract parameter choice game exists. For the two-firm case the equilibrium is unique.*

Proof. See Appendix 5.

Now to analyse numerically how equilibrium expected output as the parameters D, τ and n are varied. Again the three scenarios of the first best, moral hazard and joint profit maximisation are considered. The expected profit functions and the first-order conditions for each case are stated in Appendix 5.

Proposition 3 *As D, τ and the number of firms are increased the relative importance of moral hazard compared to joint profit maximisation decreases.*

Firstly consider the impact of D and τ on the equilibrium value of a for the two firm case. As for the case of linear demand set $R = 0$. When D is varied let $\tau = 0.1$. When τ is varied let $D = 100$. The results, in terms of total industry output, $2a$, are shown in Figure 3.

As D and τ are varied whether moral hazard or joint profit maximisation has the greater downward impact. When D and τ are small moral hazard has the greater downward impact. As D and τ grow large joint profit maximisation has the greater downward impact on expected output. This switch, again, results from the changing relative size of the agency cost against the negative revenue externality of competition. Increasing D causes the slope of the inverse demand curve to become increasingly negative:

$$\frac{\partial P(Q)}{\partial Q} = -\tau D e^{-\tau Q}$$

When D is high a given increase in output will cause a greater absolute drop in price and, hence, a larger revenue externality. When τ increases the argument is similar except that an increase in τ now leads to a greater percentage drop in price for a given percentage increase in output.[20]

Lastly the impact ofchanging the number of firms,n, on equilibrium expected output is considered. Numerical analysis equivalent to that for two firms has been performed for the cases of three and four firms. The results are shown in figures 3, 4 and 5.[21]

As n increases so does the expected industry output. The figures also show that as the n increases the values of D and τ at which collusion rather than moral hazard has the greater downward impact on output are reduced. The values of D and τ where the crossing points occur are:

Number of Firms	**Value of D**	**Value of τ**
2	3439.5	0.325
3	573.6	0.179
4	186.0	0.123

[20] Also note that as D and τ grow large the equilibrium values of a under moral hazard and in the first best converge. Whilst this result might appear odd, note that for increases in τ the equilibrium value of a falls. As a result when τ is increased the absolute size of the agency cost is reduced. Hence moral hazard has a smaller downward impact on expected output compared to the first best.

For increases in D, the result is understandable given the FOC for profit maximisation. Firm i's FOC when moral hazard is present is:

$$\frac{\partial E(\pi_i)}{\partial a_i} = \frac{D(1-\tau a_i)}{(\tau a_i+1)^3(\tau a_j+1)} - 5a_i^3 = 0$$

(See Appendix 5 for the derivation). If $a_i \geq \frac{1}{\tau}$ then this condition can never hold. As such there is an upper limit to the equilibrium value of a_i whatever the value of D.

In the first best this upper limit on a_i is reached at a lower value of D than when there is moral hazard. Under joint profit maximisation as total output is split between two firms the limit value of a is approximately half the limit value of a in the first best.

[21] In these figures total industry output, rather than output per firm, is reported and the scales on the axes vary.

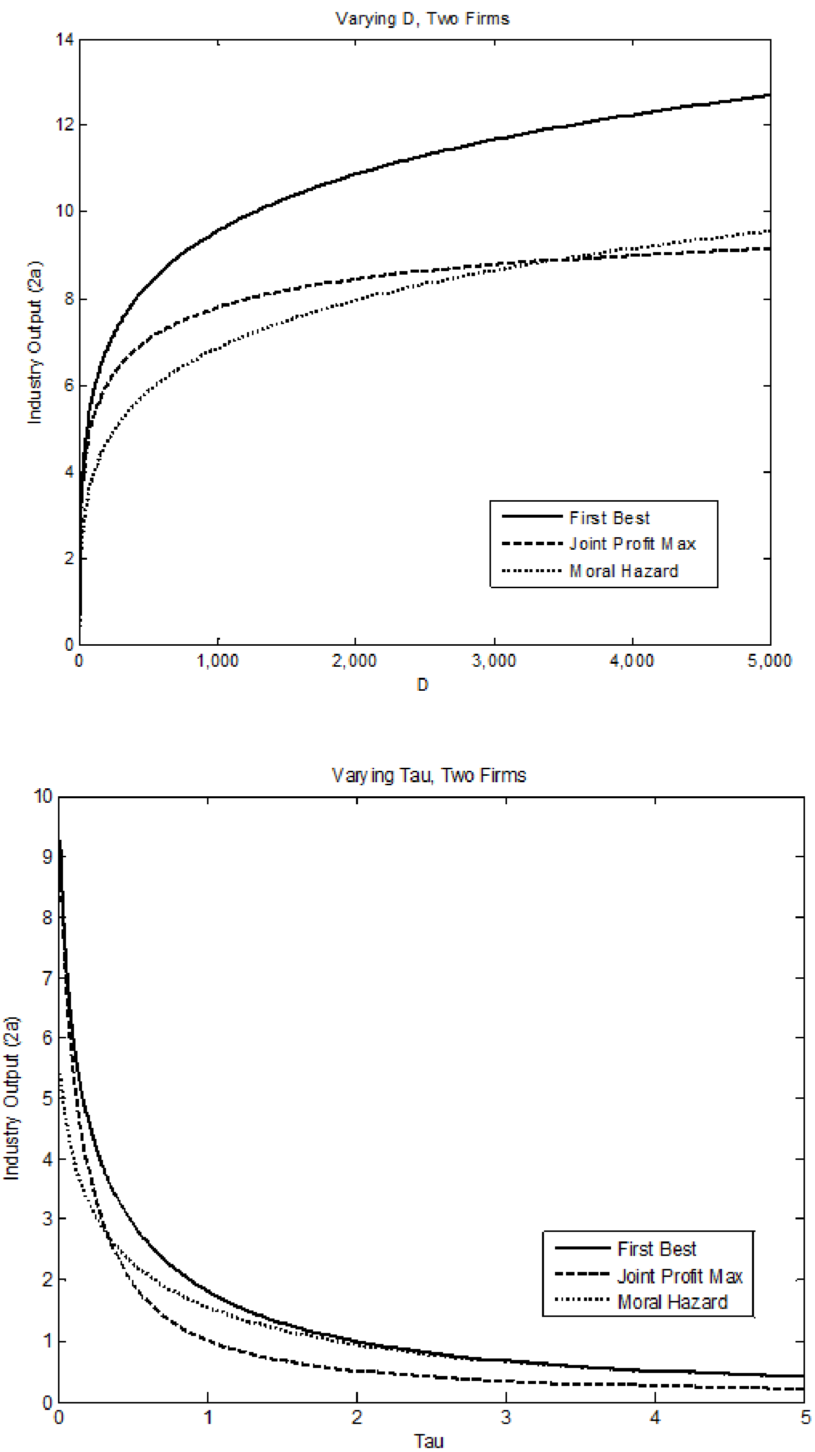
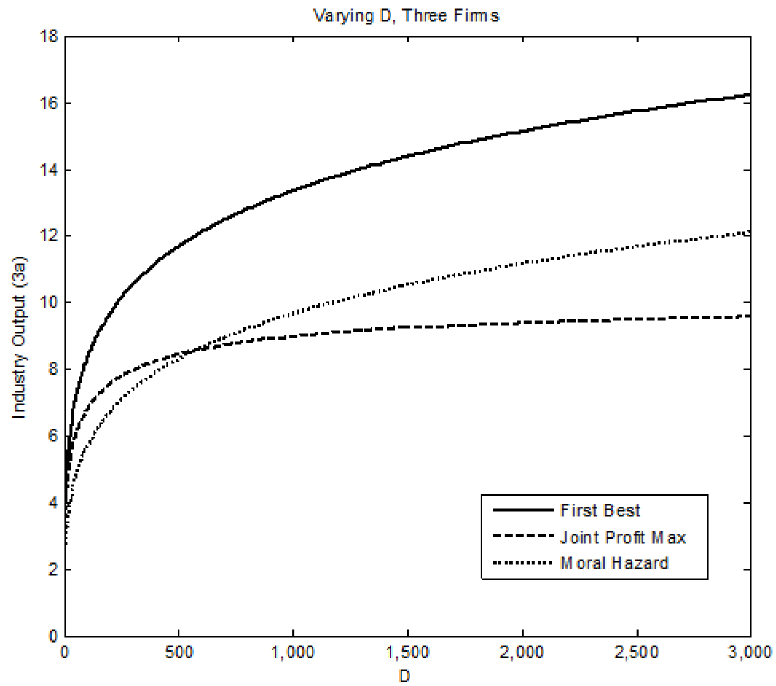


Figure 3: Expected industry output in a two firm industry as D and τ are varied



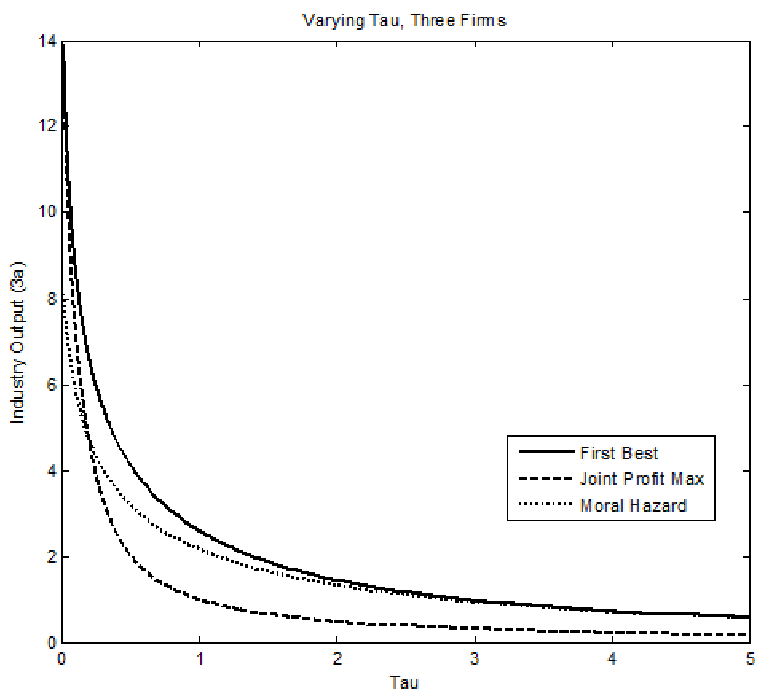
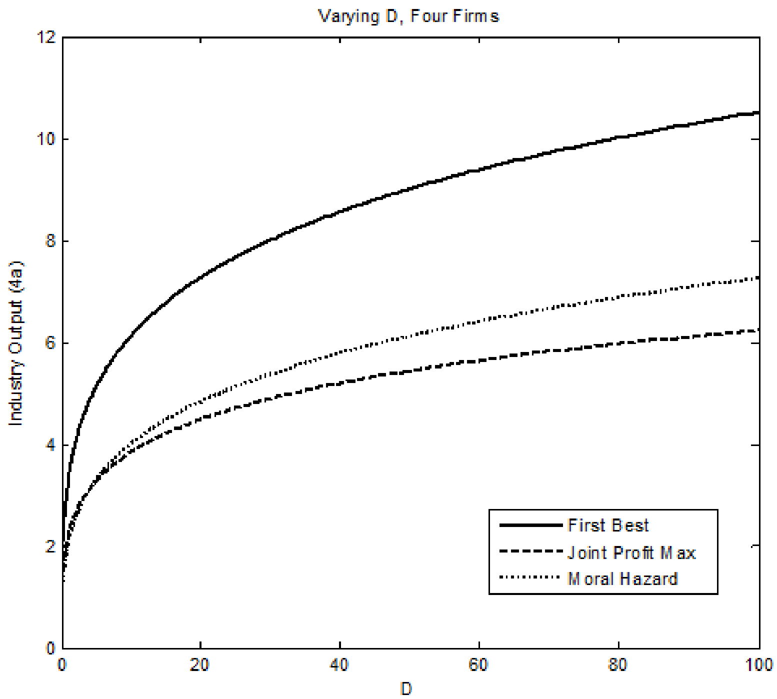


Figure 4: Expected industry output in a three-firm industry as D and τ are varied



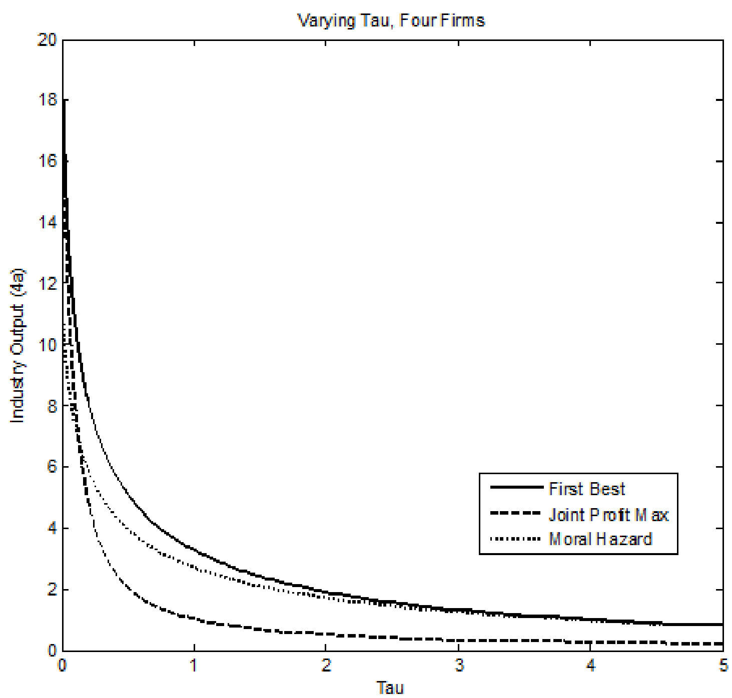


Figure 5: Expected industry output in a four-firm industry as D and τ are varied

That the values of D and τ where the crossing points occur are decreasing in n is unsurprising for two reasons. Firstly, as n increases the negative revenue externality when they do not co-operate grows. As such when joint profit maximisation internalises this externality there is a greater drop in expected industry output. Secondly, as n increases the effort each firm induces in its agent falls due to each firm having a smaller market share. Since the agency cost is increasing in effort, the cost of agency is reduced and so the downward impact of moral hazard on expected output is reduced. Intuitively it seems reasonable to suggest that this relative drop in the significance of moral hazard as n increases will be independent of the demand curve's specification.

6 Robustness and Discussion

Whilst the main results have been obtained in a specific parameterised model the basic intuitions driving them seem to apply more generally. Whether moral hazard or joint profit maximisation has a greater downward impact on expected output will depend on the relative sizes of the agency cost and the negative revenue externality between firms. The intuition that firms may not invest in agency cost reductions as they do not benefit from the resulting increases in consumer surplus also appears robust.

It should be noted that the magnitude of agency costs in this model is influenced by the assumption of each firm employing a single agent. When addressing questions relating to senior executives this assumption is probably less significant due to inherent indivisibilities in their tasks. Nevertheless incorporating multiple agents in each firm represents as important extension. With multiple workers the significance of moral hazard, as highlighted by Proposition 3, may decrease. This is because with more workers each individual worker will be required to exert less effort for the firm to produce a given level of output. Since agency costs per worker are increasing in effort as effort per workers falls so do agency costs.[22] For agency costs to remain an important determinant of market outcomes an upper limit on the number of agents employed by a firm would need to exist. To some extent the fixed costs associated with employing additional workers may limit workforce size. Indeed it would be interesting to endogenise workforce size as an initial stage game.

In a model with endogenous workforce size it would seem natural to relax the assumption of a perfectly competitive labour market. If there was a binding constraint on labour supply to increase expected output effort per worker would need

[22] However environments with multiple workers may create new agency issues. For example, if output per worker is not observable workers may free-ride on the efforts of others as in the team production literature.

to increase hence increasing the impact of agency costs. An exogenous restriction on labour supply would also limit the number of firms that could enter a market. This logic suggests that agency costs due to unobservable effort may be particularly high within professions where there are a limited number of qualified individuals, e.g. law, medicine and accountancy. In each of these professions the perception of high workloads, need to exert high effort and therefore relatively high pay seems to fit the intuition of the model.

One also needs to recognise the potential role of the exponential distribution in driving Proposition 2. For the probability density function used in section 2 increasing effort increases both the expected value and the variance of output. As a result if the principal wishes to induce additional effort the agent must be compensated both for the additional effort exerted and for being exposed to a "riskier" performance variable. To address this issue one could develop a model using LEN-framework[23] where effort reduces marginal costs. In the LEN-framework the incentive contract is linear, the agent's utility function is exponential and there is a normally distributed additive shock term. Using the normal distribution ensures the performance measure has a constant variance. However a linear contract is unlikely to be cost-minimising. Nevertheless, in practice, companies tend to use relatively simple linear incentive contracts and so developing a model reflecting this reality would have empirical merit.[24] If real-world contracts are not cost-minimising the current model would represent a lower-bound on the agency costs facing firms and the downward impact of moral hazard.[25]

On a separate note the model also demonstrates the potential role incentive contracts may play in sustaining product market collusion. In the current model, if a firm wanted to restrict output as part of a collusive arrangement they would achieve this by reducing the strength of workers' incentives in turn reducing worker effort and hence output. This hints at a neglected mechanism - using incentive contracts - which could be used to support product market collusion.[26]

Lastly if firms could commit to particular incentive contracts then agency costs themselves could be used as a strategic commitment device. For example, in this model revenue and profits are noisier performance variables than output as the former are determined by the output distributions of all firms. As a result rewards based

[23] See Holmstrom and Milgrom (1987).

[24] See Murphy (1999)

[25] That agency costs may well be higher in actual firms is also reinforced by the managerial power hypothesis of Bebchuk and Fried (2003). These authors argue that the design of incentive contracts themselves is subject to an agency problem and can be captured by management. As such contracts are likely to be biased towards enhancing agents' personal rewards rather than maximising shareholder returns.

[26] This observations fits with recent contributions by Spagnolo (2005), Chen (2008) and Aubert (2009).

on revenue or profits would be a commitment to higher agency costs and reduced expected output. However, in practice, it is questionable whether incentive contracts can have true commitment value. This is because incentive contracts are rarely fully observable and generally can be subject to re-negotiation.

7 Conclusion

The paper's key technical contribution, compared to other models combining moral hazard and competition, is deriving an optimal non-linear incentive contract where effort and output are continuous variables. Comparing the relative importance of firms co-operating over output decisions and firms' internal agency problems in determining market outcomes is also novel. The central finding from the linear demand model is that increasing market size causes moral hazard to have a greater downward impact on expected output than collusion between firms. The central finding from the exponential demand model are that increasing the number of firms or increasing the elasticity of price with respect to quantity causes the impact of moral hazard on market outcomes to become relatively less important.

The model also highlights that firms' private incentives when deciding between monitoring and incentive contracts may not always be aligned with welfare maximisation. This is notable since until the onset of the Financial Crisis concerns about managerial incentives generally focused on their potential sub-optimality from the perspective of shareholders. The non-investment in monitoring displayed here is an example of firms' incentive contracting decisions having impacts on the wider economy. However, there is some distance between recognising that these wider impacts exist and suggesting that policymakers can take concrete policy actions to address any sub-optimality.

Policy intervention regarding collusion is relatively straightforward once collusion has been proven. A standardised policy response - collusion should be stopped - can be applied since collusion is unambiguously bad. Evaluating whether the welfare gain from a reduction in agency costs outweighs the investment costs required to achieve it is a more challenging proposition and one which is context specific. Nevertheless it does seem important for policymakers and regulators to note that in specific markets private firms' solutions to agency problems may not be optimal from a welfare perspective.

Highlighting the significance of agency costs on product market outcomes also re-affirms the importance of moves to improve the oversight of managerial pay and performance. Increased reporting requirements and giving shareholders a more robust say on pay should hopefully give a greater priority to reducing agency costs.

Overall this paper emphasises the importance of considering the internal workings of firms when addressing questions in industrial organisation. The potentially significant impact of firms' internal agency problems on market outcomes seems an issue which has perhaps lacked attention in the past. Whilst it is dangerous to base policy conclusions on a single parameterised model the results do suggest that more research is warranted into this reason why all members of society, not just shareholders, have a legitimate interest in the way firms resolve agency issues.

8 Appendices

8.1 Appendix 1 - Generalised proof of existence using Frank and Quandt (1963)

8.1.1 The model

Consider a generalised n-firm setting where agents and firms can be heterogeneous. The setup of the general model is identical to that of the parameterised model unless otherwise stated. The problem is described from the perspective of firm i.

Given the n-firm setting now let $Q = \sum_{k=1}^{n} q_k$ and $A = \sum_{k=1}^{n} a_k$. Let $\phi_j = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$ be the vector of contract parameters chosen by all principals other than principal i. Continue to assume that demand is linear and continuous:

$$P(Q) = \left\{ \begin{array}{cc} B-Q, & B \geq Q \\ 0, & B < Q \end{array} \right\}$$

When $P(Q) = 0$ a finite quantity of the good is demanded. This means there is a positive real number $M < \infty$ such that $E[P(Q|A)] = 0$ for all $A \geq M$.

The probability density function for output given a specific effort level is $f(q_i|a_i)$. Assume that $f(q_i|a_i)$ is continuous and differentiable. Also assume that the cumulative distribution function $F(q_i|a_i)$ satisfies $F_{a_i}(q_i|a_i) \leq 0$ for all q_i and that for some q_i $F_{a_i}(q_i|a_i) < 0$ holds. The support of the distribution is $[\underline{q}, \infty)$ where $\underline{q} \geq 0$ thus ruling out negative quantities.[27] The distribution's support remains constant regardless of effort.

The cost of effort function, $V_i(.)$, is now assumed to be smooth, continuous, increasing and convex. The first derivative of the agent's utility function with respect to wealth, $U_i'(w)$, must be invertible.

[27] This assumption rules out, for example, the normal distribution.

8.1.2 Solving the model

To solve the model follow Holmstrom (1979) but split the principal's problem into two steps. The first step involves the principal deriving the cost-minimising incentive contract, $w_i^*(q_i)$, to induce a given level of effort a_i. For principal i this problem can be stated as:

$$\max_{w_i(q_i)} E(\pi_i) = \int_0^\infty -w_i(q_i)dF(q_i|a_i) \qquad (A1.1)$$

subject to PC:

$$\int_0^\infty U_i(w_i(q_i))dF(q_i|a_i) - V_i(a_i) \geq R_i \qquad (A1.2)$$

and IC:

$$\int_0^\infty U_i(w_i(q_i))dF_{a_i}(q_i|a_i) - V_i'(a_i) = 0 \qquad (A1.3)$$

The second step is for the principal to select the optimal contract parameter (the optimal effort level to induce) to maximise profits. This problem can be written as:

$$\max_{a_i} E(\pi_i) =$$
$$\int_0^B ... \int_0^{B-Q_{-i-j}} \int_0^{B-Q_{-i}} (B-Q)\, q_i \Pi_{k=1}^n dF(q_k|a_k) - \int_0^\infty w_i^*(q_i)dF(q_i|a_i) \qquad (A1.4)^{28}$$

where $Q = \sum_{k=1}^n q_k = q_1 + ... + q_i + q_j + ... + q_n$ and Q_{-i} is the same summation but excluding q_i.

When solving the model the objective is to obtain a pure strategy Nash equilibrium in terms of the contract parameter selected by each principal.

Step 1: Deriving the cost-minimising contract:
It is assumed that Jewitt's (1988) conditions for the validity of the first-order approach hold.

Lemma A1.1 *The cost-minimising incentive contract is:*

$$w_i^*(q_i) = U_i'^{-1}\left[\frac{1}{\lambda_i + \mu_i \frac{f_{a_i}(q_i|a_i)}{f(q_i|a_i)}}\right]$$

Given suitable assumptions on the agent's utility function and the distribution function this optimal contract will be convex. It is assumed these conditions are met.

Proof. Principal i's problem as described by (A1.1)-(A1.3) can be expressed as the following Lagrangian :

[28] If an exponential inverse demand curve is used the expected revenue function becomes $E(R_i) = \int_0^\infty ... \int_0^\infty \left(Be^{-\varepsilon Q}\right) q_i \Pi_{k=1}^n dF(q_k|a_k)$. Otherwise the proof remains the same.

$$\max_{w(q_i)} L_i = \int_0^\infty -w(q_i)dF(q_i|a_i) + \lambda_i \left[\int_0^\infty U_i(w_i(q_i))dF(q_i|a_i) - V_i(a_i) - R_i\right] +$$
$$\mu_i \left[\int_0^\infty U_i(w_i(q_i))dF_{a_i}(q_i|a_i) - V_i'(a_i)\right] \qquad (A1.5)$$

Holding a_i fixed the necessary condition for maximisation is found by taking the partial derivative of (A1.5) with respect to $w_i(q_i)$ and setting equal to zero:

$$\frac{\partial L_i}{\partial w(q_i)} = -\int_0^\infty dF(q_i|a_i) + \lambda_i \int U_i'(w_i(q_i))dF(q_i|a_i)$$
$$+\mu_i \int U_i'(w_i(q_i))dF_{a_i}(q_i|a_i) = 0$$

Realising that $dF(q_i|a_i) = f(q_i|a_i)dq_i$, dividing throughout by $U_i'(w_i(q_i))f(q_i|a_i)$ and re-arranging gives:

$$\frac{1}{U_i'(w_i(q_i))} = \lambda_i + \mu_i \frac{f_{a_i}(q_i|a_i)}{f(q_i|a_i)} \qquad (A1.6)$$

Re-arranging (A1.6) to make $w_i(q_i)$ the subject, the cost-minimising contract is:

$$w_i^*(q_i) = U_i'^{-1}\left[\frac{1}{\lambda_i + \mu_i \frac{f_{a_i}(q_i|a_i)}{f(q_i|a_i)}}\right] \qquad (A1.7)$$

The values of λ_i and μ_i can be found by inserting (A1.7) into the IC and PC and solving as two equations in two unknowns. Jewitt (1988) provides a proof for $\mu_i > 0$ which is included in Appendix 2. Given $\mu_i > 0$ the cost-minimising incentive contract is a strictly convex function of q_i if the agent's risk tolerance is sufficiently high and $\frac{f_a(q_i|a_i)}{f(q_i|a_i)}$ is a linear increasing function in q_i. The proof of this and a definition of risk tolerance, both from Basu et al (1985), are also provided in Appendix 2. ■

Step 2: Selecting the optimal value of a_i:

Assuming that Basu et al's conditions for the convexity of $w_i^*(q_i)$ hold, Jewitt's conditions on the distribution function imply that the expected wage costs, $\int_0^\infty w_i^*(q_i)dF(q_i|a_i)$ in (A1.4), will be convex. As such if the expected revenue function:

$$E(R_i) = \int_0^B ... \int_0^{B-Q_{-i,-j}} \int_0^{B-Q_{-i}} (B-Q)\, q_i \Pi_{k=1}^n dF(q_k|a_k)$$

is concave the principal's problem in (A1.4) will be concave in the contract parameter a_i. Technically all that is required for equilibrium is for $E(\pi_i)$ to be strictly quasiconcave. This ensures there is an unique profit maximising value of a_i. Adding the assumption that $E(\pi_i)$ includes a stationary point then the FOC, $\frac{\partial E(\pi_i)}{\partial a_i} = 0$, is a necessary and sufficient condition for profit maximisation.

Demonstrating existence:

This proof, adapted from Frank and Quandt (1963), involves demonstrating that the assumptions of the Kakutani fixed-point theorem hold for the current model.

Let the set $\Gamma_i = \{a_i\}$ of possible contract parameters be closed and connected for all a_i. Also define the following:

- a_i^* as the profit maximising value of a_i for the ith firm if for some fixed $\phi_j = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$, $E\left[\pi_i\left(a_i^*, \phi_j\right)\right] \geq E\left[\pi_i\left(a_i, \phi_j\right)\right]$ for all a_i.

- If an a_i^* exists, and if $a_i^* \in g_i\left(\phi_j\right)$, then g_i is the ith firm's reaction correspondence.

- $\alpha = (a_1, ..., a_n)$ and $\alpha^* = (a_1^*, ..., a_n^*)$

- The mapping $G : \{\alpha\} \rightarrow \{\alpha^*\}$ is given by the mappings $g_1 : \{\phi_1\} \rightarrow \{a_1^*\}, ..., g_n : \{\phi_n\} \rightarrow \{a_n^*\}$

An equilibrium exists in the principals' contract parameter game if it can be shown that the Kakutani Fixed Point Theorem applies to the mapping G and hence that the mapping G has a fixed point $G(a) = a$.

Lemma A1.2 *The sets $\{a_i^*\}$ are bounded for each i.*

Proof. By assumption the potential effort of each agent is bounded since $a_i \in [\underline{a}, \overline{a}]$. Also by assumption each firm is required to employ an agent and must compensate them for their effort. As such there is a lower bound to the contract parameter set at $a_i^* = \underline{a}$. Since a profit-maximising firm will always pay the agent just enough to maintain the agent's reservation utility and no more there is also an upper-bound to the contract parameter set at $a_i^* = \overline{a}$. ■

Lemma A1.3 *The sets $\{a_i^*\}$ contain at least one element.*

Proof. The functions $E\left[\pi_i\left(a_i, \phi_j\right)\right]$ are bounded by the assumptions that the expected profit function is continuous and strictly quasiconcave and by Lemma A1.2. By the continuity of both the demand function and the expected cost of the incentive contract $E\left[\pi_i\left(a_i, \phi_j\right)\right]$ must have a closed graph. As such $E\left[\pi_i\left(a_i, \phi_j\right)\right]$ has a largest element and so there exists an a_i^* such that $E\left[\pi_i\left(a_i^*, \phi_j\right)\right] \geq E\left[\pi_i\left(a_i, \phi_j\right)\right]$. ■

Lemma A1.4 *The mapping $G : \Gamma \rightarrow \Gamma$ has a fixed point.*

Proof.

1. Γ is closed by assumption and is bounded due to Lemma A1.2.

2. The mapping G maps the points $\alpha \in \Gamma$ into sets of Γ. This holds since $G(\alpha) \in \{\alpha^*\}$ by definition and $\{\alpha^*\} \subset \Gamma$.

3. The set Γ is convex. This holds because Γ is the Cartesian product of intervals.

4. The image sets $G(\Gamma)$ are convex. Assume that there are two points $\alpha^* \in G(\alpha)$ and $\alpha^{**} \in G(\alpha)$. Then for each component

$$E\left[\pi_i\left(a_i^*, \phi_j\right)\right] \geq E\left[\pi_i\left(a_i, \phi_j\right)\right];$$
$$E\left[\pi_i\left(a_i^{**}, \phi_j\right)\right] \geq E\left[\pi_i\left(a_i, \phi_j\right)\right];$$

and choosing γ, $0 < \gamma < 1$,

$$E\left[\pi_i\left(a_i,\phi_j\right)\right] \leq \min\left\{E\left[\pi_i\left(a_i^*,\phi_j\right)\right], E\left[\pi_i\left(a_i^{**},\phi_j\right)\right]\right\} < E\left[\pi_i\left(\gamma a_i^* + (1-\gamma)\, a_i^{**},\phi_j\right)\right],$$

where the latter inequality holds due to the strict quasiconcavity of $E\left[\pi_i\left(a_i,\phi_j\right)\right]$. As a result the entire line segment between a_i^* and a_i^{**} is in the image $G(\alpha)$.

5. G is upper semi-continuous. For this to be true it is necessary that for every sequence $\alpha^\rho \to \alpha^0$, such that $\alpha^{*\rho} \in G(\alpha^\rho)$ and $\alpha^{*\rho} \to \alpha^{*0}$, it is the case that $\alpha^{*0} \in G(\alpha^0)$. Follow a proof by contradiction. If the contrary held then

$$E\left[\pi_i\left(a_i^{*0},\phi_j^0\right)\right] < E\left[\pi_i\left(g_i\left(\phi_j^0\right),\phi_j^0\right)\right]$$

for some i. As a result there would be points $\alpha^{*\rho}$ arbitrarily close to α^{*0} for which $\alpha^{*\rho} \notin G(\alpha^\rho)$, which is a contradiction.

So by the Kakutani Fixed Point Theorem there exists α such that $G(\alpha) = \alpha$. As such the mapping G has a fixed point. ■

Hence existence of equilibrium in a generalised n-firm setting with linear demand has been demonstrated if suitable assumptions hold.

8.2 Appendix 2 - Proofs by other authors[29]

8.2.1 Conditions for the validity of the first-order approach - Jewitt (1988)

Jewitt states that the conditions for the first-order approach to be valid are:

(i) $\int_{-\infty}^{q} F(q|a)dq$ is non-increasing convex in a for each value of q,

(ii) $\int qdF(q|a)dq$ is non-decreasing concave in a,

(iii) $\frac{f_a(q|a)}{f(q|a)}$ is non-decreasing concave in q for each value of a, and

(iv) the utility of the agent is a concave increasing function of the observable variables, i.e. $\omega(z) = U\left(U'^{-1}\left(\frac{1}{z}\right)\right)$, where $z > 0$, is concave.

These conditions essentially ensure the agent's problem is concave. To understand how they do this let $w(q)$ solve the first-order condition:

$$\int_0^\infty U\left(w\left(q\right)\right) dF_a\left(q|a\right) - V'\left(a\right) = 0 \qquad (A2.1)$$

Using the fact that $\mu > 0$[30], condition (iii) and that for the cost-minimising contract:

[29] For ease of notation the i subscript is dropped in this Appendix.

[30] See section 8.2.2 for Jewitt's proof that $\mu > 0$.

$$\frac{1}{U'(w(q))} = \lambda + \mu\frac{f_a(q|a)}{f(q|a)} \qquad (A2.2)^{31}$$

it is possible to say that $\frac{1}{U'(w(q))}$ is non-decreasing concave. Condition (iv) is the requirement that $U(w(q))$ is a concave transformation of $\frac{1}{U'(w(q))}$. Since the class of non-decreasing concave functions is closed under composition it means that $U(w(q))$ is non-decreasing concave in q. Lastly it is necessary to demonstrate that the transformation φ to φ^* defined by:

$$\varphi^*(a) = \int_0^\infty \varphi(q)dF(q,a)$$

preserves concavity. If this is the case then the agent's problem will be concave and the solution to the first-order condition will be a global maximum. Conditions (i) and (ii) are necessary for this concavity preserving property to hold. ■

8.2.2 Proof that $\mu > 0$ - Jewitt (1988)

Lemma A2.1 *If U is an increasing concave function and $V'(a) > 0$ then any μ satisfying (A2.1) and (A2.2) is positive.*

Proof. Re-arranging (A2.2) gives:

$$f_a(q|a) = \frac{1}{\mu}\left(\frac{1}{U'(w(q))} - \lambda\right) f(q|a)$$

Substituting this into (A2.1) and re-arranging gives:

$$\int U(w(q))\left(\frac{1}{U'(w(q))} - \lambda\right) f(q|a)dq = \mu V'(a) \qquad (A2.3)$$

Since $\int f(q|a)dq = 1$ it must be the case that $\int f_a(q|a)dq = 0$ and in turn:

$$E\left(\frac{f_a(q|a)}{f(q|a)}\right) = 0$$

So considering (A2.2) it is possible to write:

$$E\left(\frac{1}{U'(w(q))}\right) = \lambda \qquad (A2.4)$$

Now consider the LHS of (A2.3)

$$E\left[\int U(w(q))\left(\frac{1}{U'(w(q))} - \lambda\right) f(q|a)dq\right] = \int \frac{U(w(q))}{U'(w(q))} f(q|a)dq - \lambda\int U(w(q))f(q|a)dq$$

[31] For the derivation of this condition see the proof of Lemma A1.1.

Given (A2.4) this expression gives the covariance between $U(w(q))$ and $\frac{1}{U'(w(q))}$. As such:

$$Cov\left(U(w(q)), \tfrac{1}{U'(w(q))}\right) = \mu V'(a)$$

Since both U and $\frac{1}{U'}$ are monotone increasing functions they must have a non-negative covariance and since $V'(a) > 0$ by assumption it follows that $\mu \geq 0$. $\mu = 0$ can be ruled out since it would imply that $w(q)$ would be a constant by re-arrangement of (A2.2). Having $w\,(q)$ as a constant would violate the incentive compatibility constraint and hence $\mu > 0$ must hold. ■

8.2.3 Conditions for the cost-minimising incentive contract to be convex - Basu et al (1985)

The following lemma is taken, with changed notation, from Appendix A of Basu et al (1985).

Lemma A2.2 *If* $\frac{f_a(q|a)}{f(q|a)}$ *is a linear function of* q *over some interval then the cost-minimising contract,* $w(q)$*, is a strictly convex function of* q *over that interval if the rate of change of risk tolerance,* $T'(w)$*, exceeds one.*

Proof. Since $\frac{f_a(q|a)}{f(q|a)}$ is linear in q then (A2.2) can be written as:

$$\tfrac{1}{U'(w(q))} = A + Bq$$

Taking the second derivative with respect to q of both sides of this expression gives:

$$\tfrac{U'''(w)[w'(q)]^2 + U''(w)w''(q)}{[U'(w)]^2} - \tfrac{2[U''(w)]^2[w'(q)]^2}{[U'(w)]^3} = 0$$

Dividing throughout by $\frac{U''(w)}{[U'(w)]^3}$ and simplifying leads to:

$$U'(w)w''(q) = \left[2U''(w) - \tfrac{U'(w)U'''(w)}{U''(w)}\right][w'(q)]^2$$

Since $U'(w) > 0$ and $U''(w) < 0$

$$sign\,[w''(q)] = sign\left[\tfrac{U'(w)U'''(w)}{[U''(w)]^2} - 2\right] \qquad (A2.5)$$

The Arrow-Pratt measure of absolute measure of risk aversion is given by:

$$R_a(w) = \tfrac{-U''(w)}{U'(w)}$$

Since $R_a(w)$ measures risk aversion its inverse, $T(w)$, given by:

$$T(w) = \frac{1}{R_a(w)} = \frac{-U'(w)}{U''(w)},$$

measures risk tolerance. Differentiating $T(w)$ by w gives:

$$T'(w) = \frac{U'(w)U'''(w)}{[U''(w)]^2} - 1$$

Combining this with (A2.5) leads to:

$$sign\,[w''(q)] = sign\,[T'(w) - 1]$$

and so it follows that $w''(q) > 0$ if and only if $T'(w) > 1$. ■

A power utility function of the form $U = \frac{1}{\delta}w^{\delta}$ where $0 < \delta < 1$ displays the necessary risk tolerance.

8.3 Appendix 3 - Proof of FOCs in the parameterised model

8.3.1 The First Best (observable and verifiable effort)

Lemma A3.1 *When effort is observable and verifiable the FOC for a symmetric equilibrium is:*

$$\frac{1}{a^2}e^{-\frac{B}{a}}\left(\frac{1}{3}B^3 + \frac{3}{2}B^2a + 4Ba^2 + 6a^3\right) + B - 5a - a^3 = 0$$

Proof. Solve the principal's problem in two steps.

Step 1: Find the cost-minimising contract to induce a given effort a_i from agent i. Since effort is observable and verifiable principal i can induce agent i to exert an effort a_i by offering a forcing contract. This contract will involve a payment just satisfying agent i's participation constraint if the effort a_i is exerted and a payment of zero if any other effort level is observed. As such there is no incentive compatibility constraint in the principal's Lagrangian:

$$\begin{aligned} \max_{w_i(q_i)} L_i &= -\int_0^\infty w_i(q_i)\frac{1}{a_i}e^{-\frac{q_i}{a_i}}dq_i \\ &+\lambda_i\left[\int_0^\infty 2\left(w_i(q_i)\right)^{\frac{1}{2}}\frac{1}{a_i}e^{-\frac{q_i}{a_i}}dq_i - a_i^2 - R\right] \qquad (A3.1) \end{aligned}$$

Fix the value of a_i. Now take the partial derivative of (A3.1) with respect to $w_i(q_i)$ and set equal to zero to give:

$$\frac{\partial L_i}{\partial w_i(q_i)} = -\int_0^\infty \frac{1}{a_i}e^{-\frac{q_i}{a_i}}dq_i + \lambda_i\int_0^\infty \left(w_i(q_i)\right)^{-\frac{1}{2}}\frac{1}{a_i}e^{-\frac{q_i}{a_i}}dq_i = 0 \qquad (A3.2)$$

Re-arranging and simplifying (A3.2) gives:

$$w_i^*(q_i) = \lambda_i^2 \qquad (A3.3)$$

i.e. the cost-minimising contract is a constant wage. Assume the principal's participation constraint binds with equality. Inserting (A3.3) into the participation constraint and re-arranging gives:

$$\lambda_i = \frac{a_i^2+R}{2} \qquad (A3.4)$$

Inserting (A3.4) back into the expression for $w_i^*(q_i)$ gives the cost-minimising contract for inducing an effort level a_i as:

$$w_i^*(q_i) = \left(\frac{a_i^2+R}{2}\right)^2 \qquad (A3.5)$$

Step 2: Now principal i's problem is to select the optimal value of a_i to write in the contract $w_i^*(q_i)$. Note the expected revenue function, $E(R_i)$, is unchanged from section 3. Setting $R = 0$ the expected cost of inducing the effort a_i, i.e. the expected cost of $w_i^*(q_i)$, is simply:

$$\left(\frac{a_i^2}{2}\right)^2$$

As a result principal i's second step problem is:

$$\max_{a_i} E(\pi_i) =$$
$$a_i a_j^3 \frac{e^{-\frac{B}{a_j}}}{(a_i-a_j)^2} + a_i^2 e^{-\frac{B}{a_i}} \frac{Ba_i - Ba_j + 2a_i^2 - 3a_i a_j}{(a_i-a_j)^2} + (B - 2a_i - a_j)\, a_i - \left(\frac{a_i^2}{2}\right)^2 \qquad (A3.6)$$

The FOC of this problem is:

$$\frac{\partial E(\pi_i)}{\partial a_i} =$$
$$-a_j^3 \frac{e^{-\frac{B}{a_j}}}{(a_i-a_j)^3}(a_i+a_j) + \frac{e^{-\frac{B}{a_i}}}{(a_i-a_j)^3}\begin{pmatrix} B^2a_i^2 - 2B_i^2 a_i a_j + B^2 a_j^2 + 3Ba_i^3 \\ -8Ba_i^2 a_j + 5Ba_i a_j^2 + 4a_i^4 - 11a_i^3 a_j + 9a_i^2 a_j^2 \end{pmatrix}$$
$$+B - 4a_i - a_j - a_i^3 = 0$$

Appealing to the problem's symmetry, combining the first two terms and applying l'hôpital's rule three times the FOC for a symmetric equilibrium is:

$$\frac{1}{a^2} e^{-\frac{B}{a}} \left(\frac{1}{3}B^3 + \frac{3}{2}B^2 a + 4Ba^2 + 6a^3\right) + B - 5a - a^3 = 0 \qquad (A3.7)$$

■

8.3.2 Firms Maximise Joint Profits (collusion)

Lemma A3.2 *When firms maximise joint profits, and effort is observable and verifiable, the FOC for a symmetric equilibrium is:*

$$\tfrac{1}{2a^2}e^{-\frac{B}{a}}\left(B^3+4B^2a+10Ba^2+12a^3\right)+B-6a-a^3=0$$

Proof. **Step 1:** Since effort is observable and verifiable the cost-minimising contracts $w_i^*(q_i)$ and $w_j^*(q_j)$ take the same form as in the proof of Lemma A3.1:

$$w_i^*(q_i)=\left(\tfrac{a_i^2}{2}\right)^2 \text{ and } w_j^*(q_j)=\left(\tfrac{a_j^2}{2}\right)^2$$

Step 2: To find the optimal values for a_i and a_j it is easiest to consider the firms as a single monopolist with two agents. The problem facing the combined firm is:

$$\max_{a_i,a_j} E\left(\pi_{i+j}\right)=E\left(R_i\right)+E\left(R_j\right)-\left(\tfrac{a_i^2}{2}\right)^2-\left(\tfrac{a_j^2}{2}\right)^2$$

where

$$E\left(R_i\right)+E\left(R_j\right)=\tfrac{1}{a_i-a_j}\left(a_i^2e^{-\frac{B}{a_i}}\left(B+2a_i\right)-a_j^2e^{-\frac{B}{a_j}}\left(B+2a_j\right)\right)+a_i\left(B-2a_i-a_j\right)+a_j\left(B-a_i-2a_j\right) \qquad (A3.8)$$

There are two FOCs for this problem, $\frac{\partial E(\pi_{i+j})}{\partial a_i}=0$ and $\frac{\partial E(\pi_{i+j})}{\partial a_j}=0$. The FOC with respect to a_i is:

$$\frac{\partial E(\pi_{i+j})}{\partial a_i}=\frac{1}{(a_i-a_j)^2}\left(\begin{array}{c} e^{-\frac{B}{a_i}}\left(4a_i^3-6a_i^2a_j+3Ba_i^2+B^2a_i-B^2a_j-4Ba_ia_j\right)\\ -a_j^2e^{-\frac{B}{a_j}}\left(B+2a_j\right)\end{array}\right)+B-4a_i-2a_j-a_i^3=0$$

Applying symmetry and using l'hôpital's rule three times the FOC for maximisation of joint profits becomes:

$$\tfrac{1}{2a^2}e^{-\frac{B}{a}}\left(B^3+4B^2a+10Ba^2+12a^3\right)+B-6a-a^3=0 \qquad (A3.9)$$

■

8.4 Appendix 4 - Proof of Proposition 3

Proposition A4.1 *The ranges of investment costs where a social planner would invest in a perfect monitoring technology but two competing firms would not are:*

Market Size (B)	**Investment Cost Range**
50	$37.5 < C < 45.6$
100	$113.9 < C < 125.6$
200	$308.8 < C < 333.9$

Proof. The proof is split into two parts. The first involves finding the ranges of investment costs where different Nash equilibria occur in the investment subgame. The second is to find the ranges of investment costs where a social planner maximising total surplus would choose to invest/not invest. Comparing these cost ranges then leads to Proposition A4.1. The proof described is for $B = 100$. For other values of B the procedure is qualitatively identical but involves different numerical values.[32]

Lemma A4.1 *In the initial investment subgame:*

- *if* $C < 111.0$ *Invest is the strictly dominant strategy and the Nash equilibrium is:* $(Invest, Invest)$
- *if* $111.0 < C < 113.9$ *there are two Nash equilibria:* $(Invest, NotInvest)$ *and* $(NotInvest, Invest)$
- *if* $C > 113.9$ *NotInvest is the strictly dominant strategy and the Nash equilibrium is:* $(NotInvest, NotInvest)$

Proof. To determine the equilibrium decisions of each firm the pay-off matrix for the two principals must be formed. This involves to identifying the optimal contract parameters, a_i^* and a_j^*, which maximise the expected profits of each firm given different combinations of investment decisions. If firm i chooses $NotInvest$ a_i is selected to maximise (9) and if firm i chooses $Invest$ a_i is selected to maximise (A3.6).[33] The values of a_i^* and a_j^* for each of the decision pairs, when $B = 100$, are shown below:

Decision Pairs	a_i^* , a_j^* to 3d.p.
$(Invest, Invest)$	$4.283, 4.283$
$(Invest, NotInvest)$	$4.312, 2.575$
$(NotInvest, Invest)$	$2.575, 4.312$
$(NotInvest, NotInvest)$	$2.592, 2.592$

[32] The full workings for $B = 50$ and $B = 200$ are available on request.

[33] When only one firm invests the problem facing the firms is no longer symmetric and so $a_i^* = a_j^*$ no longer holds. The two first-order conditions with respect to a_i and a_j therefore are solved as a system of two equations in two unknowns.

Inserting a_i^* and a_j^* back into the firms' profit functions allows the firms' profits gross of investment costs to be obtained. Subtracting the investment cost, C, when a firm chooses $Invest$ gives the following pay-off matrix:

		Principal j	
		$Invest$	$NotInvest$
Principal i	$Invest$	$289.1 - C, 289.1 - C$	$296.5 - C, 178.2$
	$NotInvest$	$178.2, 296.5 - C$	$182.6, 182.6$

By comparing the pay-offs between different decision pairs the Nash equilibria of the investment subgame can be characterised for different values of C. This gives Lemma A4.1. ■

Now consider the decision of a social planner maximising total surplus. The social planner has three options: invest in neither firm, invest in one firm or invest in both firms. For now assume that the social planner can only make different investment decisions compared to profit maximising firms.[34] As a result the values of a_i^* and a_j^* stated in the proof of Lemma A4.1 are still used in the calculations below.

Lemma A4.2 *A social planner maximising total surplus will:*

- *invest in the monitoring technology for both firms if* $C < 122.3$
- *invest in the monitoring technology for one firm if* $122.3 < C < 125.6$
- *not invest if* $C > 125.6$

Total surplus gross of investment costs is given by:

$$E\left(TS_g\right) = \int_0^B \int_0^{B-q_j} \left(B\left(q_i + q_j\right) - \frac{\left(q_i+q_j\right)^2}{2}\right) \frac{1}{a_i} e^{-\frac{q_i}{a_i}} \frac{1}{a_j} e^{-\frac{q_j}{a_j}} dq_i dq_j + \frac{(B)^2}{2}\left(\frac{1}{a_i - a_j}\left(a_i e^{-\frac{B}{a_i}} - a_j e^{-\frac{B}{a_j}}\right)\right) - E\left(w_i\left(q_i\right)\right) - E\left(w_j\left(q_j\right)\right) \qquad (A4.1)$$

where

$$B\left(q_i + q_j\right) - \frac{\left(q_i+q_j\right)^2}{2}$$

is the area under the inverse demand curve when $q_i + q_j \leq B$ and

$$\frac{(B)^2}{2}\left(\frac{1}{a_i - a_j}\left(a_i e^{-\frac{B}{a_i}} - a_j e^{-\frac{B}{a_j}}\right)\right)$$

[34] The possibility of the social planner directly setting effort/production levels is ruled out. From a policy perspective this seems a reasonable distinction to make. In the US and EU investment subsidies are fairly common whereas policymakers micro-managing firms' operational decisions is much rarer.

is the area under the inverse demand curve when $q_i + q_j > B$ multiplied by the probability of $q_i + q_j > B$.

The probability that $q_i + q_j > B$ is simply:

$$P(q_i + q_j > B) = 1 - P\left(q_i + q_j \leq B\right)$$

where:

$$P\left(q_i + q_j \leq B\right) = \int_0^B \int_0^{B-q_j} \frac{1}{a_i} e^{-\frac{q_i}{a_i}} \frac{1}{a_j} e^{-\frac{q_j}{a_j}} dq_i dq_j = 1 - \frac{1}{a_i - a_j}\left(a_i e^{-\frac{B}{a_i}} - a_j e^{-\frac{B}{a_j}}\right)$$

After completing the necessary integration and substituting in the relevant values of a_i^* and a_j^* the following values of total surplus net of investment costs are obtained:

Invest in both firms	Invest in one firm	Don't Invest
$633.3 - 2C$	$511.0 - C$	385.4

By comparing these values Lemma A4.2 is obtained. ■

Comparing the cost ranges for different investment decisions in Lemmas A4.1 and A4.2 then leads to Proposition A4.1. ■

8.5 Appendix 5 - Proofs involving exponential inverse demand

8.5.1 Expected profits are strictly quasiconcave

To demonstrate that the expected profit function, $E(\pi_i)$, is strictly quasiconcave firstly the expected revenue function needs to be derived for the n-firm case.

Lemma A5.1 *The expected revenue function for firm i in an n-firm setting is given by:*

$$E\left(R_i\right) = \frac{Da_i}{(\tau a_i+1)^2 \Pi_{k\neq i}^n (\tau a_k+1)}$$

Proof. The expected revenue function for firm i in an n-firm setting can be written as:

$$E\left(R_i\right) = \int_0^\infty ... \int_0^\infty \int_0^\infty De^{-\tau Q} q_i \Pi_{k=1}^n \frac{1}{a_k} e^{-\frac{qk}{a_k}} dq_k$$

where firm i is just a particular firm between 1 and n. For the following explanation it is helpful to re-write the above expression as:

$$E\left(R_i\right)=\int_0^{\infty}...\int_0^{\infty}\int_0^{\infty}De^{-\tau(q_i+Q_{-i})}q_i\frac{1}{a_i}e^{-\frac{q_i}{a_i}}dq_i\Pi_{k\neq i}^{n}\frac{1}{a_k}e^{-\frac{qk}{a_k}}dq_k \qquad (A5.1)$$

where $Q_{-i}=\left(\sum_{k=1}^{n}q_k\right)-q_i$.

Now it is necessary to integrate with respect to each dq_k beginning with dq_i. Since the probability density functions for the quantities produced by firms other than i are independent of q_i they can be moved outside of the integral with respect to dq_i and can be dealt with subsequently. As such the first integral to consider is:

$$\int_0^{\infty}De^{-\tau(q_i+Q_{-i})}q_i\frac{1}{a_i}e^{-\frac{q_i}{a_i}}dq_i=\int_0^{\infty}\frac{D}{a_i}e^{\frac{-(\tau a_i+1)q_i-\tau a_iQ_{-i}}{a_i}}q_idq_i=\frac{Da_i}{(\tau a_i+1)^2}e^{-\tau Q_{-i}}$$

After this first integration $E\left(R_i\right)$ can be written as:

$$E\left(R_i\right)=\int_0^{\infty}...\int_0^{\infty}\frac{Da_i}{(\tau a_i+1)^2}e^{-\tau(q_j+Q_{-i-j})}\frac{1}{a_j}e^{-\frac{q_j}{a_j}}dq_j\Pi_{k\neq i,j}^{n}\frac{1}{a_k}e^{-\frac{qk}{a_k}}dq_k$$

where $Q_{-i-j}=\left(\sum_{k=1}^{n}q_k\right)-q_i-q_j$. Considering the integral with respect to dq_j gives:

$$\int_0^{\infty}\frac{Da_i}{(\tau a_i+1)^2}e^{-\tau(q_j+Q_{-i-j})}\frac{1}{a_j}e^{-\frac{q_j}{a_j}}dq_j=\frac{Da_i}{(\tau a_i+1)^2}\int_0^{\infty}\frac{1}{a_j}e^{-\left(\frac{\left(\tau a_j+1\right)q_j+\tau a_jQ_{-i-j}}{a_j}\right)}=\frac{Da_i}{(\tau a_i+1)^2(\tau a_j+1)}e^{-\tau Q_{-i-j}} \qquad (A5.2)$$

As a result $E\left(R_i\right)$ becomes:

$$E\left(R_i\right)=\int_0^{\infty}...\int_0^{\infty}\frac{Da_i}{(\tau a_i+1)^2(\tau a_j+1)}e^{-\tau Q_{-i-j}}\Pi_{k\neq i,j}^{n}\frac{1}{a_k}e^{-\frac{qk}{a_k}}dq_k$$

Integrating with respect to each subsequent dq_k is qualitively identical to the operation performed when integrating with respect to dq_j. For example, if the next integration is with respect to dq_l the resulting expression, equivalent to $(A5.2)$, is:

$$\frac{Da_i}{(\tau a_i+1)^2(\tau a_j+1)(\tau a_l+1)}e^{-\tau Q_{-i-j-l}}$$

Once all the integration procedures have been performed the expected revenue function can be expressed as:

$$E\left(R_i\right)=\frac{Da_i}{(\tau a_i+1)^2\Pi_{k\neq i}^{n}(\tau a_k+1)} \qquad (A5.3)$$

■

Having found $E\left(R_i\right)$, and since the structure of the agency relationship remains identical to that in Section 2, the expected profit function for firm i when agency costs are present is:

$$E\left(\pi_i\right) = \frac{Da_i}{(\tau a_i+1)^2 \Pi_{k\neq i}^n (\tau a_k+1)} - \frac{5}{4}a_i^4 \qquad (A5.4)$$

To demonstrate that this function is strictly quasiconcave in a_i note that if a function is strictly increasing before it is strictly decreasing it is strictly quasiconcave. Hence for $E\left(\pi_i\right)$ to be strictly quasiconcave it is necessary to show that: (i) $\frac{\partial E(\pi_i)}{\partial a_i}$ changes sign only once over the range $a_i \in [\underline{a}, \overline{a}]$, (ii) this sign change is from positive to negative, and (iii) if $\frac{\partial E(\pi_i)}{\partial a_i} = 0$ it occurs at only one point.

The expression for $\frac{\partial E(\pi_i)}{\partial a_i}$ is:

$$\frac{\partial E(\pi_i)}{\partial a_i} = \frac{D(1-\tau a_i)}{(\tau a_i+1)^3 \Pi_{k\neq i}^n (\tau a_k+1)} - 5a_i^3 \qquad (A5.5)$$

As long as $\underline{a} > 0$ is sufficiently small $\frac{\partial E(\pi_i)}{\partial a_i}$ will start as a positive value.[35] By inspection both $\frac{D(1-\tau a_i)}{(\tau a_i+1)^3 \Pi_{k\neq i}^n (\tau a_k+1)}$ and $-5a_i^3$ are strictly decreasing in a_i and so $\frac{\partial E(\pi_i)}{\partial a_i}$ is a strictly decreasing function in a_i. Assuming $\overline{a}$ is sufficiently large not to act as a constraint then as a_i grows large $\frac{\partial E(\pi_i)}{\partial a_i}$ will turn negative and stay negative. Since $\frac{\partial E(\pi_i)}{\partial a_i}$ is a strictly decreasing function it also means that this function will cross the horizontal axis only once implying that $\frac{\partial E(\pi_i)}{\partial a_i} = 0$. Hence it has been shown that the expected profit function is strictly quasiconcave over the range $a_i \in [\underline{a}, \overline{a}]$.

Since $E\left(\pi_i\right)$ is strictly quasiconcave then by Theorem 1 an equilibrium will exist in the firms' contract parameter choice game.

8.5.2 Uniqueness of the two firm equilibrium

In the case of two firms the first-order condition for firm i to be profit-maximising is:

$$\frac{\partial E(\pi_i)}{\partial a_i} = \frac{D(1-\tau a_i)}{(\tau a_i+1)^3 (\tau a_j+1)} - 5a_i^3 = 0 \qquad (A5.6)$$

Re-arranging this condition it is possible to write a_j as an explicit function of a_i:

$$a_j = \frac{D(1-\tau a_i)}{5\tau a_i^3 (\tau a_i+1)^3} - \frac{1}{\tau} \qquad (A5.7)^{36}$$

By symmetry there will be an equivalent equation for a_i in terms of a_j:

[35] $\frac{\partial E(\pi_i)}{\partial a_i}$ will be positive whenever:

$$D > \frac{5a_i^3(\tau a_i+1)^3 \Pi_{k\neq i}^n (\tau a_k+1)}{(1-\tau a_i)}$$

Note that as a_i tends to zero the right-hand side of this inequality tends to zero. Hence the requirement for $\frac{\partial E(\pi_i)}{\partial a_i}$ to be positive at $\underline{a}$ tends to the condition $D > 0$ as $\underline{a}$ becomes small.

[36] Note that (A5.7) and (A5.8) are not the reaction functions for firm's i and j.

$$a_i = \frac{D(1-\tau a_j)}{5\tau a_j^3(\tau a_j+1)^3} - \frac{1}{\tau} \qquad (A5.8)$$

To demonstrate an unique equilibrium exists a basic geometric argument can be used. The aim is to prove that the lines described by (A5.7) and (A5.8) cross once, and only once, in (a_i, a_j)-space. Firstly note that both equations are continuous. Also note that because $a_i, a_j \in [\underline{a}, \overline{a}]$ both (A5.7) and (A5.8) are bounded and closed. Lastly recall the starting assumption that $\underline{a}$ and $\overline{a}$ are sufficiently far apart never to impinge on the equilibrium outcome. Combining continuity, boundness, closedness and the symmetry of the problem the lines represented by (A5.7) and (A5.8) must cross at least once.

The next stage is demonstrate that the lines represented by (A5.7) and (A5.8) cross only once. This involves demonstrating two things: (i) both (A5.7) and (A5.8) are decreasing convex over the range $[\underline{a}, \overline{a}]$, and (ii) (A5.7) and (A5.8) do not coincide.

Lemma A5.2 *The functions (A5.7) and (A5.8) are both decreasing convex over the range of* $a_i, a_j \in [\underline{a}, \overline{a}]$.

Proof. Due to symmetry if one of (A5.7) and (A5.8) is proven to be decreasing convex the other function will also be decreasing convex. Consider (A5.7) only. The function is decreasing as long as:

$$\frac{\partial a_j}{\partial a_i} = -\frac{1}{5\tau}\frac{D}{a_i^4(\tau a_i+1)^4}\left(-5\tau^2 a_i^2 + 4\tau a_i + 3\right) < 0$$

which will be the case if

$$-5\tau^2 a_i^2 + 4\tau a_i + 3 > 0 \qquad (A5.9)$$

By inspection of (A5.7) a_j is gauranteed to be negative once $a_i \geq \frac{1}{\tau}$. Setting the left-hand side of (A5.9) equal to zero and finding the roots of the resulting equation it is possible to say that when $\frac{4-2\sqrt{19}}{10\tau} < a_i < \frac{4+2\sqrt{19}}{10\tau}$ it implies $\frac{\partial a_j}{\partial a_i} < 0$. Since $\frac{4-2\sqrt{19}}{10\tau} < 0$ and $\underline{a} > 0$ we only need to check that $a_i < \frac{4+2\sqrt{19}}{10\tau}$ holds for the range of a_i being considered. Consider $a_i = \frac{1}{\tau}$. Since $\frac{1}{\tau} < \frac{4+2\sqrt{19}}{10\tau}$ it means that for any positive value of a_j such that $a_i \geq \underline{a} > 0$ $\frac{\partial a_j}{\partial a_i} < 0$ holds. Hence (A5.7) is decreasing in a_i and (A5.8) is decreasing in a_j for the region of (a_i, a_j)-space being considered.

For the function represented by the right-hand side of (A5.7) to be convex requires:

$$\frac{\partial^2 a_j}{\partial a_i^2} = \frac{6}{5\tau}\frac{D}{a_i^5(\tau a_i+1)^5}\left(-5\tau^3 a_i^3 + 3\tau^2 a_i^2 + 6\tau a_i + 2\right) > 0$$

which will hold if:

$$-5\tau^3 a_i^3 + 3\tau^2 a_i^2 + 6\tau a_i + 2 > 0 \qquad (A5.10)$$

holds. Again using the assumption $a_j \geq \underline{a} > 0$ it is sufficient to demonstrate that (A5.7) is convex when a_i satisfies $\underline{a} \leq a_i \leq \frac{1}{\tau}$. The argument for $\frac{\partial^2 a_j}{\partial a_i^2} > 0$ is that (A5.10) is a cubic and so has two stationary points. It can be shown that the two stationary points of (A5.10) are at $\frac{6-2\sqrt{99}}{30\tau}$ and $\frac{6+2\sqrt{99}}{30\tau}$. Only the second stationary point occurs for a positive value of a_i. By inspection the a_i^3 term has a negative coefficient and hence when $|a_i|$ is large (A5.10) will be decreasing. Hence when $a_i \geq \underline{a}$ (A5.10) will be quasiconcave. Also note that when $a_i = \underline{a}$ for $\underline{a}$ small enough that value of (A5.10) is $2 > 0$. When $a_i = \frac{1}{\tau}$ the value of (A5.10) is $6 > 0$. Since (A5.10) is postive at $a_i = \underline{a}$ and at $a_i = \frac{1}{\tau}$ and is also quasiconcave between these two points means that for the relevant range of a_i (A5.10) is positive. In turn this means that $\frac{\partial^2 a_j}{\partial a_i^2} > 0$ for the relevant region of (a_i, a_j)-space and so (A5.7) is decreasing convex as required. ■

That (A5.7) and (A5.8) are both decreasing convex when combined with the other assumptions implies that they must either coincide or cross only once.

Lemma A5.3 *The lines represented by (A5.7) and (A5.8) do not coincide.*

Proof. Follow a proof by contradiction. Firstly consider the case where $a_i \neq a_j$. Assume the lines do coincide. Then it must be the case that when either $a_i = \underline{a}$ or $a_j = \underline{a}$ both (A5.7) and (A5.8) must hold.

Consider the case where $a_j = \underline{a}$ and denote the value of a_i which solves (A5.7) and (A5.8) as $\widehat{a}_i$. Then (A5.7) and (A5.8) can be re-written as:

$$5\tau\widehat{a}_i^3 (\tau\widehat{a}_i + 1)^3 (\tau\underline{a} + 1) = D\tau (1 - \tau\widehat{a}_i) \qquad (A5.11)$$

and

$$5\tau\underline{a}^3 (\tau\underline{a} + 1)^3 (\tau\widehat{a}_i + 1) = D\tau (1 - \tau\underline{a}) \qquad (A5.12)$$

Assume $\widehat{a}_i > \underline{a}$, this implies that the right-hand side of (A5.11) has a smaller value than the right-hand side of (A5.12). However $\widehat{a}_i > \underline{a}$ also implies that the left-hand side of (A5.11) has a higher value than the left-hand side of (A5.12). As a result there cannot be a value of $\widehat{a}_i$ which satisfies both (A5.11) and (A5.12). Hence there is a contradiction and (A5.7) and (A5.8) do not coincide when $a_i > \underline{a}$. As $a_j = \underline{a}$ the case of $a_i < a_j$ does not need to be considered. If $a_j > \underline{a}$ making $a_i < a_j$ the logic of the proof would still hold. Also, by symmetry, the same arguments hold when we hold a_i fixed and vary a_j.

Now consider the case where $a_i = a_j = a$ and again assume both lines coincide. The only line that can satisfy these two conditions is a straight line decreasing at an angle of 45 degrees. The second derivative of such a line must be zero. However from Lemma A5.2 it is known that for plausible values of $a_i, a_j > \underline{a}$: $\frac{\partial^2 a_j}{\partial a_i^2} > 0$ and $\frac{\partial^2 a_i}{\partial a_j^2} > 0$ from (A5.7) and (A5.8) respectively. Hence there is a contradiction. (A5.7) and (A5.8) cannot be satisfied whilst $a_i = a_j = a$ and both lines coincide.$\blacksquare$

Since (A5.7) and (A5.8) do not coincide it must be the case that they cross only once. Hence an equilibrium exists and it must be unique. The same form of argument also demonstrates an unique equilibrium exists when the case of observable and verifiable effort is considered.

8.5.3 FOCs for two firms

Moral Hazard:

In the subsection above (A5.6) gives the FOC for profit maximisation by firm i. Re-arranging this condition and applying symmetry it is straightforward to show that the equilibrium value of $a_i = a_j = a$ must satisfy:

$$D - 5a^7\tau^4 - 20a^6\tau^3 - 30a^5\tau^2 - 20a^4\tau - 5a^3 - Da\tau = 0$$

First Best:

When effort is observable and verifiable the expected profit function for firm i is:

$$E(\pi_i) = \frac{Da_i}{(\tau a_i+1)^2(\tau a_j+1)} - \frac{1}{4}a_i^4$$

and the FOC is:

$$\frac{\partial E(\pi_i)}{\partial a_i} = \frac{D(1-\tau a_i)}{(\tau a_i+1)^3(\tau a_j+1)} - a_i^3 = 0$$

Re-arranging and applying symmetry the equation that the equilibrium value of a must satisfy is:

$$D - a^7\tau^4 - 4a^6\tau^3 - 6a^5\tau^2 - 4a^4\tau - a^3 - Da\tau = 0$$

Joint Profit Maximisation:

The profit maximisation problem faced by the firms is:

$$\max_{a_i,a_j} E(\pi_{i+j}) = \frac{Da_i}{(\tau a_i+1)^2(\tau a_j+1)} + \frac{Da_j}{(\tau a_j+1)^2(\tau a_i+1)} - \frac{1}{4}a_i^4 - \frac{1}{4}a_j^4$$

The two FOCs are:

$$\frac{\partial E(\pi_{i+j})}{\partial a_i} = \frac{D\left(1-2a_ia_j\tau^2-a_i\tau\right)}{(\tau a_i+1)^3(\tau a_j+1)^2} - a_i^3 = 0$$
$$\frac{\partial E(\pi_{i+j})}{\partial a_j} = \frac{D\left(1-2a_ia_j\tau^2-a_j\tau\right)}{(\tau a_j+1)^3(\tau a_i+1)^2} - a_j^3 = 0$$

Applying symmetry and re-arranging both of these conditions reduce to:

$$D - a^8\tau^5 - 5a^7\tau^4 - 10a^6\tau^3 - 10a^5\tau^2 - 5a^4\tau - a^3 - 2Da^2\tau^2 - Da\tau = 0$$ [37]

8.5.4 FOCs when three and four firms

The FOCs used to find the equilibrium values of a in the three and four firm cases are provided below. They are obtained using a procedure equivalent to that for the two firm case.

Three Firms:
First Best:

$$D - a^8\tau^5 - 5a^7\tau^4 - 10a^6\tau^3 - 10a^5\tau^2 - 5a^4\tau - a^3 - Da\tau = 0$$

Moral Hazard:

$$D - 5a^8\tau^5 - 25a^7\tau^4 - 50a^6\tau^3 - 50a^5\tau^2 - 25a^4\tau - 5a^3 - Da\tau = 0$$

Joint Profit Maximisation:

$$D - a^{10}\tau^7 - 7a^9\tau^6 - 21a^8\tau^5 - 35a^7\tau^4 - 35a^6$$
$$\tau^3 - 21a^5\tau^2 - 7a^4\tau - 3Da^3\tau^3 - a^3 - 5Da^2\tau^2 - Da\tau = 0$$

Four Firms:
First Best:

$$D - a^9\tau^6 - 6a^8\tau^5 - 15a^7\tau^4 - 20a^6\tau^3 - 15a^5\tau^2 - 6a^4\tau - a^3 - Da\tau = 0$$

Moral Hazard:

$$D - 5a^9\tau^6 - 30a^8\tau^5 - 75a^7\tau^4 - 100a^6\tau^3 - 75a^5\tau^2 - 30a^4\tau - 5a^3 - Da\tau = 0$$

Joint Profit Maximisation:

$$D - a^{12}\tau^9 - 9a^{11}\tau^8 - 36a^{10}\tau^7 - 84a^9\tau^6 - 126a^8\tau^5 - 126a^7\tau^4 - 84a^6\tau^3 - 36a^5\tau^2 - 4Da^4$$
$$\tau^4 - 9a^4\tau - 11Da^3\tau^3 - a^3 - 9Da^2\tau^2 - Da\tau = 0$$

[37] Due to the functional forms of $\frac{\partial E(\pi_{i+j})}{\partial a_i}$ and $\frac{\partial E(\pi_{i+j})}{\partial a_j}$, for now, assume that the problem is strictly quasiconcave.

9 References

- Aghion, P., Dewatripont, M. and Rey, P. (1999), "Agency Costs, Firm Behaviour and the Nature of Competition", CEPR Discussion Paper No.2130
- Aubert, C. (2009), "Managerial effort incentives and market collusion", Toulouse School of Economics Working Paper series 09-127
- Basu, A.K., Lal, R., Srinivasan, V. and Staelin R. (1985), "Salesforce Compensation Plans: An Agency Theoretic Perspective", Marketing Science, 4(4), pg 267-291
- Bebchuk, L.A. and Fried, J.M. (2003), "Executive Compensation as an Agency Problem", Journal of Economic Perspectives, 17(3), pg 71-92
- Bhardwaj, P. (2001), "Delegating Pricing Decisions", Marketing Science, 20(2), pg 143-169
- Bonatti, A. (2003), "The Strategic Choice of Contractual Policies", Rivista Di Politica Economica, Nov-Dec 2003, pg 167-190
- Chen, Z. (2008), "Cartel Organization and Antitrust Enforcement", CCP Working Paper 08-21, University of East Anglia, Norwich, England
- Fershtman, C. and Judd, K.L. (1987), "Equilibrium Incentives in Oligopoly", American Economic Review, 77(5), pg 927-940
- Frank Jr, C.R. and Quandt, R.E. (1963), "On the Existence of Cournot Equilibrium", International Economic Review, 4(1), pg 92-96
- Fumas, V.S. (1992), "Relative performance evaluation of management: The effects on industrial competition and risk sharing", International Journal of Industrial Organisation, 10(3), pg 473-489
- Gal-Or, E. (1993), "Internal organization and managerial compensation in oligopoly", International Journal of Industrial Organisation, 11(2), pg 157-183
- Gal-Or, E. (1997), "Multiprincipal Agency Relationships as Implied by Product Market Competition", Journal of Economics and Management Strategy, 6(2), pg 235-256
- Grossman, S.J. and Hart, O.D. (1983), "An Analysis of the Principal-Agent Problem", Econometrica, 51(1), pg 7-45
- Hart, O.D. (1983), "The Market Mechanism as an Incentive Scheme", Bell Journal of Economics, 14(2), pg 366-382

- Hermalin, B.E. (1994), "Heterogeneity in Organizational Form: Why Otherwise Identical Firms Choose Different Incentives for their Managers", RAND Journal of Economics, 25(4), pg 518-537

- Holden, R.T. (2008), "Does Competition Make Firms More Efficient?", Working Paper, MIT, Cambridge, Massachusetts

- Holmstrom, B. (1979), "Moral Hazard and Observability", The Bell Journal of Economics, 10(1), pg 74-91

- Holmstrom, B. and Milgrom, P. (1987), "Aggregation and linearity in the provision of intertemporal incentives", Econometrica, 55(2), pg303-328

- Jensen, M.C. and Meckling, W.H. (1976), "Theory of the firm: Managerial behaviour, agency costs and ownership structure", Journal of Financial Economics, 3(4), pg 305-360

- Jewitt, I. (1988), "Justifying the First-Order Approach to Principal-Agent Problems", Econometrica, 56(5), pg 1177-1190

- Leibenstein, H. (1966), "Allocative Efficiency vs "X-Inefficiency"", American Economic Review, 56(3), pg 392-415

- Mirrlees, J.A. (1999), "The Theory of Moral Hazard and Unobservable Behaviour", Review of Economic Studies, 66(1), pg 3-21

- Mirrlees, J.A. (1976), "The Optimal Structure of Incentives and Authority within an Organization", Bell Journal of Economics, 7(1), pg 105-131

- Mishra, B.K. and Prasad, A. (2005), "Delegating Pricing Decisions in Competitive Markets with Symmetric and Asymmetric Information", Marketing Science, 24(3), pg 490-497

- Murphy, K.J. (1999), "Executive Compensation", in "Handbook of Labour Economics, Volume 3" eds. Ashenfelter, O. and Card, D., pg 2485-2563, Elsevier Science BV, North-Holland

- Plehn-Dujowich, J.M. and Serfes, K. (2010), "Strategic Managerial Compensation Arising From Product Market Competition", Working Paper, Temple University, Philadelphia

- Raith, M. (2003), "Competition, Risk and Managerial Incentives", American Economic Review, 93(4), pg 1425-1436

- Scharfstein, D. (1988), "Product Market Competition and Managerial Slack", RAND Journal of Economics, 19(1), pg 147-155

- Schleifer, A. and Vishny (1986), "Large Shareholders and Corporate Control", Journal of Political Economy, 94(3), pg 461-488
- Schmidt, K.M. (1997), "Managerial Incentives and Product Market Competition", Review of Economic Studies, 64(2), pg 191-213
- Sklivas, S.D. (1987), "The Strategic Choice of Managerial Incentives", RAND Journal of Economics, 18(3), pg 452-458
- Spagnolo, G. (2005), "Managerial incentives and collusive behaviour", European Economic Review, 49(6), pg 1501-1523
- Stiglitz, J.E. (1974), "Incentives and Risk Sharing in Sharecropping", Review of Economic Studies, 41(2), pg 219-255
- Szidarovsky, F. and Yakowitz, S. (1977), "A New Proof of the Existence and Uniqueness of the Cournot Equilibrium", International Economic Review, 18(3), pg 787-789
- Theilen, B. (2008), "Market Competition and Lower Tier Incentives", Working Paper, Universitat Rovira i Virgili, Spain